Harvard Obs. 5205
APR - 7 1919
Univ. of Mich.

VOLUME XLIX NUMBER 2

THE ASTROPHYSICAL JOURNAL

AN INTERNATIONAL REVIEW OF SPECTROSCOPY AND ASTRONOMICAL PHYSICS

EDITED BY

GEORGE E. HALE
Mount Wilson Solar Observatory of the Carnegie Institution of Washington

EDWIN B. FROST
Yerkes Observatory of the University of Chicago

HENRY G. GALE
Major, U.S.A. in France

MARCH 1919

THEORY OF IMPERFECT GRATINGS . . . . . C. M. Sparrow 65

STUDIES BASED ON THE COLORS AND MAGNITUDES IN STELLAR CLUSTERS. X Harlow Shapley 95

AERIAL PHOTOMETRY . . . . . M. Luckiesh 108

REVIEWS:
The Binary Stars, Robert Grant Aitken (G. Van Biesbroeck) 121.

EDWARD CHARLES PICKERING 136

THE UNIVERSITY OF CHICAGO PRESS
CHICAGO, ILLINOIS, U.S.A.

THE CAMBRIDGE UNIVERSITY PRESS, London and Edinburgh
THE MARUZEN-KABUSHIKI-KAISHA, Tokyo, Osaka, Kyoto, Fukuoka, Sendai
THE MISSION BOOK COMPANY, Shanghai

March 1919 Vol. XLIX, No. 2

THE ASTROPHYSICAL JOURNAL

AN INTERNATIONAL REVIEW OF SPECTROSCOPY AND ASTRONOMICAL PHYSICS

EDITED BY

GEORGE E. HALE
Mount Wilson Solar Observatory of the Carnegie Institution of Washington

EDWIN B. FROST
Yerkes Observatory of the University of Chicago

HENRY G. GALE
Major, U.S.A. in France

WITH THE COLLABORATION OF

EDWIN S. AMES, Johns Hopkins University
ANASTASIOS BELOPOLSKY, Observatoire de Poulkovo
WILLIAM W. CAMPBELL, Lick Observatory
EDWARD KASNER, Northwestern University
CHARLES BARRY, Université de Marseille
CHARLES S. HASTINGS, Yale University
HENDRICUS KAYSER, Universität Bonn
ALBERT A. MICHELSON, University of Chicago
HUGH F. NEWALL, Cambridge University
ERNEST F. NICHOLS, Yale University
ALFRED PEROT, Paris
†EDWARD C. PICKERING, Harvard College Observatory
ANNIBALE RICCO, Osservatorio di Catania
CARL RUNGE, Universität Göttingen
ARTHUR SCHUSTER, The University, Manchester
FRANK SCHLESINGER, Allegheny Observatory

†Died Feb. 3, 1919.

The Astrophysical Journal is published by the University of Chicago at the University of Chicago Press, 5750 Ellis Avenue, Chicago, Illinois, during each month except February and August. ¶The subscription price is $5.00 a year; the price of single copies is 65 cents. Orders for service of less than a half-year will be charged at the single-copy rate. ¶Postage is prepaid by the publishers on all orders from the United States, Mexico, Cuba, Porto Rico, Panama Canal Zone, Republic of Panama, Hawaiian Islands, Philippine Islands, Guam, Samoan Islands, Shanghai. ¶Postage is charged extra as follows: For Canada, 38 cents on annual subscriptions (total $5.38), on single copies, 3 cents (total 68 cents); for all other countries in the Postal Union, 62 cents on annual subscriptions (total $5.62), on single copies, 11 cents (total 76 cents). ¶Patrons are requested to make all remittances payable to The University of Chicago Press in postal or express money orders or bank drafts.

The following are authorized to quote the prices indicated:

For the British Empire: The Cambridge University Press, Fetter Lane, London, E.C. 4. Yearly subscriptions, including postage, 25s. each; single copies, including postage, 3s. 3d. each.

For Japan and Korea: The Maruzen-Kabushiki-Kaisha, 11 to 16 Nihonbashi Tori Sanchome, Tokyo, Japan. Yearly subscriptions, including postage, Yen 11.25 each; single copies, including postage, Yen 1.20 each.

For China: The Mission Book Company, 18 Peking Road, Shanghai. Yearly subscriptions, $5.00; single copies, 65 cents, or their equivalents in Chinese money. Postage extra, if mailed direct outside of Shanghai, on yearly subscriptions 62 cents, on single copies 11 cents.

Claims for missing numbers should be made within the month following the regular month of publication. The publishers expect to supply missing numbers free only when losses have been sustained in transit, and when the reserve stock will permit. Deliveries are not guaranteed in the war zones of Europe, but in other countries where transportation is disturbed by conditions incident to war.

Business communications should be addressed to The University of Chicago Press, Chicago, Illinois. Communications for the editors and manuscripts should be addressed to the Editors of The Astrophysical Journal, Yerkes Observatory, Williams Bay, Wisconsin.

The cable address of the Editorial Office is "Arcturus, Chicago."

This publication is indexed in the *Readers' Guide Supplement*, New York, N.Y.

Entered as second-class matter, January 17, 1895, at the Post-office at Chicago, Ill., under the act of March 3, 1879. Acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, authorized on

THE
ASTROPHYSICAL JOURNAL
AN INTERNATIONAL REVIEW OF SPECTROSCOPY
AND ASTRONOMICAL PHYSICS

VOLUME XLIX MARCH 1919 NUMBER 2

THEORY OF IMPERFECT GRATINGS

By C. M. SPARROW

A complete bibliography of the theory of imperfect gratings would make a long list of papers by many writers. Of these writers, four perhaps stand out from the others. The classical researches of Rowland[1] and of Rayleigh,[2] employing what may be called the conventional methods of analysis, form the basis of most of the work that has been done. Besides these, Cornu[3] has developed a method of peculiar elegance for investigating the focal peculiarities of gratings, and Michelson[4] a method which is specially adapted to the study of periodic errors of spacing. In the present paper an attempt is made to present the whole subject afresh from the standpoint of the so-called vector method. This method, used so brilliantly by Cornu in treating elementary diffraction phenomena, was first applied to the perfect grating by A. L. Kimball.[5] It is a method of great power for some problems, enabling us to avoid the use of much complicated analysis, particularly in those problems in which

[1] *Phil. Mag.* (5), **35**, 397, 1893; *Astronomy and Astrophysics*, **12**, 129, 1893; *Physical Papers*, p. 525.

[2] *Coll. Papers*, I, 415; III, 47.

[3] *Comptes Rendus*, **116**, 1215, 1893. [4] *Astrophysical Journal*, **18**, 298, 1903.

[5] *Phil. Mag.* (6), **6**, 30, 1903; see also Wood, *Physical Optics*, p. 203, for a simple exposition.

qualitative insight is sought rather than a numerical solution. Where the numerical result is desired, the graphic solution obtained from the geometric representation is often the shortest road; in other cases we can pass at once from the vector diagram to its analytic equivalent, the geometric intuition serving as a guide in keeping the analysis in its simplest form.

The first part of the paper considers systematically the various kinds of defects which are likely to occur in practice. Many well-known results will appear in the new garb; many others, not so well known, are nevertheless contained by implication in the already published work of others, but have hitherto failed to receive the emphasis which their importance demanded. In order to increase this emphasis, curves showing the forms of diffraction patterns have been included, and numerical data which have been obtained by ordinary processes of calculation. The second part of the paper applies some of the results obtained in the first part to an examination of the limitations of the method of coincidences as a means of measuring wave-length. It contains, in addition, a theory of the defective Fabry and Perot interferometer with consequences relating to this method of measuring wave-lengths.

In the discussion that follows we shall consider the grating to have $2N$ grooves, and shall in general denote by n the number of a particular groove, counted from the center of the grating, n being positive or negative according as the groove is on one side or the other of the center. It will be necessary to recall briefly the main features of the vector treatment of the perfect plane grating. Let Δs_n and ϕ_n be the amplitude and phase (with respect to the central groove) of the disturbance due to the nth groove. If we represent this disturbance by a vector of length Δs_n and polar angle ϕ_n the vector sum of all such vectors will represent the amplitude and phase of the resultant disturbance. Intensities are obtained by squaring the amplitudes. For the perfect grating, Δs_n is independent of n, and

$$\phi_n - \phi_{n-1} \equiv \Delta\phi_n$$

is constant. The order in which the vectors may be added is completely arbitrary, but if we add them in the order of increasing n,

the vector polygon obtained is equilateral and equiangular and is therefore inscriptible in a circular arc. Let us denote by θ the angle of diffraction, and by θ_0, θ_1, ... the angles corresponding to the central image, first order spectrum, etc. For $\theta=\theta_0$, $\Delta\phi=0$, and the vector polygon is a straight line of length $2N\Delta s$. As we pass away from this direction, $\Delta\phi$ increases, and the vector polygon curls up, the length of the broken line remaining constant, until, when $\Delta\phi=\pi/N$, the polygon closes completely: this is the direction of the first minimum. As we go still farther, the polygon laps over itself on the circumscribing circle, giving the first subsidiary maximum when it has gone about one and a half times round, the second minimum when it has gone twice round, and so on. (The well-known formula for the intensity,

$$J=\frac{J_0 \sin^2 N\Delta\phi}{\sin^2 \Delta\frac{\phi}{2}}$$

is obtainable at once from the vector diagram without the use of trigonometric identities.) The brightness of the maxima steadily decreases until $\Delta\phi=\pi$. From this point on, an increasing $\Delta\phi$ is equivalent to a negative phase-difference of decreasing absolute value (i.e., $\Delta\phi$ is replaced by $\Delta\phi-2\pi$). The vector polygon gradually straightens out, going through the previous series of changes in the reverse order, until in the direction θ_1 it becomes perfectly straight again. Between this and the second order the same changes are repeated, and so on.

At the pth minimum on either side of a principal maximum,

$$\Delta\phi=p\pi/n.$$

So long as this is small, the polygon will approach very closely to the circumscribing circle, and can be replaced by it. As N is in practice quite large and the values of p which interest us are usually small, we are led to consider an infinite number N of infinitesimal vectors ds, of phase-difference $d\phi$. Instead of the vector polygon we have a vector curve. But

$$\frac{ds}{d\phi}=\rho,$$

the radius of curvature of the curve. In the case of the perfect grating just considered ds and $d\phi$ are independent of n, so that ρ is a constant. In general ds and $d\phi$ are functions of n, so that

$$\rho = f(n).$$

Thus f is expressible as a function of s or ϕ, or both, giving us the intrinsic equation of the vector curve. The form of f is determined by the nature of the defects of the grating, and the properties of such gratings are completely expressed by the properties of the family of vector curves; just as the properties of the perfect grating are expressed by those of a family of circular arcs of constant length.

The defects of gratings fall into two fundamentally distinct classes according as they involve a variation of $d\phi$ or ds. We shall call them phase errors and amplitude errors, respectively. The phase errors are practically the most important; they have also received the fullest treatment by analytic methods. The amplitude errors which actually occur are usually such as to produce only minor defects in the performance of the grating, but we should include under this head certain "defects" which may be artificially introduced in order to accomplish some particular end. We will begin with a discussion of phase errors, and will thus assume that ds is independent of n.

In the direction θ_k of a principal maximum of the kth order, the phase-difference between successive disturbances should be constant and equal to $2k\pi$, provided the grating is perfect and the instrument is in adjustment. That this is not the case is due in practice to one or more of several causes: (1) the spacing of the lines is not uniform; (2) the surface is not accurately plane; (3) the telescope and collimator are not in adjustment.[1] The possibility arises of several of these causes being present at once and producing errors which more or less compensate; this is in fact actually the case with the concave grating. The essential results of the theory are all included in a treatment of errors due to non-uniform spa-

[1] Crookedness of the lines is a fourth defect, which does not readily lend itself to treatment by the vector method. See the papers of Rayleigh referred to above for a discussion of the consequences of this class of error.

cing.[1] We shall assume, therefore, unless otherwise stated, that the surface of the grating is perfectly plane, and that telescope and collimator are in adjustment, but that the distance between the $(n-1)$th and nth grooves is a function of n, $f(n)$, the nature of which will be further specified.

The defects of spacing of actual ruled gratings may be divided into four classes: (1) Errors of "run," a small progressive change of spacing caused by variation of temperature during the ruling, etc. (2) A sudden "jump" or discontinuity in the ruling, the grating interval being constant except at a single point. Michelson has pointed out the possibility of an error of this kind; the actual frequency of its occurrence is a matter of speculation. (3) A periodic variation in the spacing, usually of the pitch of the screw, producing "ghosts." (4) Accidental variations of spacing, inherent in the imperfections of the ruling engine.

Each of these kinds of error, except the last, has its appropriate form of $f(n)$. Instead of $f(n)$ it will be convenient to represent the grating interval by $a+f(n)$, where a is the interval at the center. The phase-difference between successive disturbances is then

$$\frac{2\pi \sin \theta}{\lambda}[a+f(n)].$$

From this we must subtract $2k\pi$ for the spectrum of the kth order, in order to get $d\phi$. We are concerned only with the phenomena near a principal maximum. We may therefore put

$$\frac{2\pi a \sin \theta}{\lambda} - 2k\pi = d\phi_0,$$

where $d\phi_0$ is thus the excess of phase-difference over $2k\pi$ for two grooves at the center of the grating. Thus

$$d\phi = d\phi_0 + \frac{2\pi \sin \theta}{\lambda} f(n).$$

But we have also

$$\frac{2\pi a \sin \theta k}{\lambda} = 2k\pi,$$

[1] Conversely, any distribution of the grooves has its parallel in the aberrations of an optical instrument, the treatment given here being equally applicable.

and since we are only concerned with values of θ near θ_k, and the variation of spacing is also small, we may replace θ by θ_k, giving

$$d\phi = d\phi_0 + \frac{2k\pi}{a} f(n).$$

The length ds of the elementary vectors being constant, n is proportional to s, the length of arc of the vector curve, measured from the center (say $n = Qs$). Hence dividing by ds, we have finally

$$c = c_0 + kF(s) \tag{1}$$

$$F(s) \equiv Af(Qs)$$

where c is the curvature of the vector curve at the point s and c_0 the curvature at the center of the curve. c_0 is that parameter of the family of curves which varies with θ, the diffraction angle. The effect of the error is thus to superpose upon the curvature c_0 a curvature which varies from point to point of the curve, but which is independent of c_0 and hence of θ. The fact that errors of spacing produce no effect on the form of the central image is an immediate corollary of this.

Certain general properties of these curves may now be noted. A given length of arc of the curve corresponds to a definite portion of the ruled surface. The value of c_0 determines the point of the diffraction pattern, being proportional to the phase of this point. Draw the osculating circle (curvature c_0) at the center of the vector curve, and take an arc $2s_0$ equal in length to the vector curve. This circular arc is the vector curve for a perfect grating of spacing a; it corresponds to the same diffraction angle θ as the actual vector curve. Now the angle between the tangents at the ends of this arc is $2c_0s_0$; this then is the extreme variation in phase between the end grooves of the grating of spacing a. If we choose s_0 equal to unity we see that c_0 is half the phase of the corresponding point of the diffraction pattern, the zero of phase being taken as that corresponding to the *direction of the principal maximum of an ideal grating of spacing a*. While this direction furnishes a convenient zero it is not the proper standard. We should rather take as our zero the phase corresponding to the direction of a principal maximum of an ideal grating with the *same mean spacing as the actual grating*. A

perfect grating with the same mean spacing would have the same total distance between the first and last grooves as the actual grating. In the direction of a principal maximum the phase of these two grooves is the same, hence in the actual grating the vector curve which corresponds to the proper zero of phase has parallel tangents at its two ends. It has seemed necessary to go into this point in some detail, because it has proved a source of confusion in the work of some other investigators. It is characteristic of the vector method that it fixes attention on the law of spacing, by means of the variation of curvature, rather than on the total "aberration," which is indicated by the variation of the slope of the curve to the polar axis.

Errors of run.—As we may assume these to be small, we may write

$$c = c_0 + \beta s + \gamma s^2 + \delta s^3 + \ldots \qquad (2)$$

We shall speak of the successive terms as representing linear, quadratic, etc., errors, these corresponding to the aberrations of the second, third, etc., degree of ordinary optics. There is a fundamental difference between the terms of odd degree and those of even degree. For the errors of odd degree, the middle of the vector curve $c_0 = 0$ is a center of symmetry, the curvatures on each side of this point being equal and opposite. The tangents at the two ends of this curve being parallel, the mean spacing is the same as the spacing at the center. If we superpose now a uniform curvature c_0, the shape of the curve will be independent of the sign of c_0; the diffraction pattern is therefore symmetrical with respect to the true direction of the principal maximum. Thus for errors of odd degree, *the spectral images are symmetrical and in their true position.* As we pass away from the center of the pattern, the point of inflexion of the curve moves toward one end; the curvature of one half is thus always greater than that of the other half, and the curve will not close completely. The minima of intensity will thus not be "clean" but will always have some illumination. For small errors, the amplitude of this "error of closure" of the vector curve will be proportional to the error, and hence to the order of the spectrum. The intensity of illumination at the minimum will thus be proportional to the square of the order.

The simplest case of an error of odd degree is represented by

$$c=c_0+\beta s$$

or

$$\phi=c_0 s+\tfrac{1}{2}\beta s^2.$$

The vector curves are Cornu spirals, as may be seen by putting

$$s=s'-\frac{c_0}{\beta}.$$

This furnishes a simple proof of the well-known fact that an error of this type simply changes the focus of the grating. Just

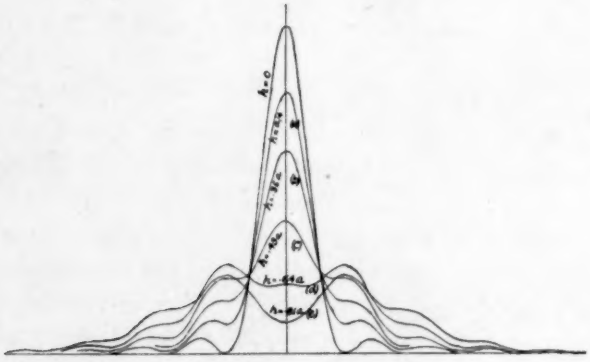

FIG. 1

as the form of the spectral images of a perfect grating in exact focus is identical, for large N, with that of an aperture of the width of the ruled surface, so the diffraction pattern of such a grating out of focus can be found in the usual way by the Cornu spiral. From the linearity of the curvature formula (2) it follows that small errors may be considered as producing their effects independently, their curvatures being simply superposed. Therefore by a suitable change of focus the effect of a linear error may be removed, the curvature of the vector curve becoming uniform, and the diffraction pattern that of a perfect grating.

A few curves are given (Fig. 1) showing the form of these diffraction images. It will be noted that they embody the general features of the effect of the errors of odd degree mentioned above. They

may be interpreted either as due to a linear error of spacing or as the extra-focal images of a perfect grating. The quantity h denotes the maximum displacement of a groove from its true position.

Cubic error.—The next case in order of simplicity is where

$$c = c_0 + \delta s^3$$

or

$$\phi = c_0 s + \tfrac{1}{4}\delta s^4.$$

The general form of the vector curve resembles the Cornu spiral, but the convolutions converge more rapidly to the limit points. We cannot obtain all our intensities from a single curve, as with the Cornu spiral, but must use the entire family of curves of which c_0 is the variable parameter. The length of spiral used depends on the magnitude of the error. A few of the spirals are shown in Fig. 2 (solid lines). They are so drawn that for $c_0 = 0$ the spiral turns through one right angle in the length 10 (marked by small circles). The diffraction pattern corresponding to this error is shown in the solid curve (*a*)

FIG. 2

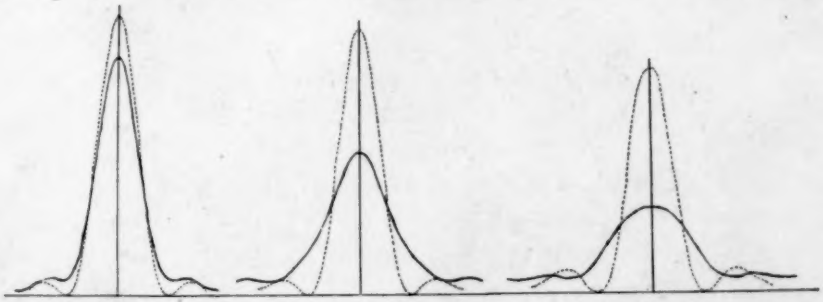

FIG. 3

of Fig. 3. If this curve is compared with (*a*) of Fig. 1, which corresponds to the same value of h, it will be seen that the distortion of the spectral image is less for the cubic error. This is to

be expected, as the central portions of the spirals of Fig. 2 are straighter than those of the Cornu spiral. The length 12 of the spirals, indicated by crosses, corresponds to

$$h=1\cdot 2^4\cdot a/4=0.5184a.$$

The corresponding diffraction pattern is shown in the solid curve (*b*) of Fig. 3. The curve (*c*) of Fig. 1 corresponds to a somewhat smaller value of h, but the distortion is greater. The whole length of spiral corresponds to

$$h=1\cdot 4^4\cdot a/4=0.9604a;$$

the diffraction pattern is shown in (*c*) Fig. 3. This error amounts to a maximum shift of groove of practically a whole grating interval, but there is still some definition left. A linear error of the same size destroys it almost completely.

The question now arises: How far can the effect of this cubic error be eliminated by focusing? The effect of focusing is to superpose a linear error of opposite sign, so that the vector spiral is given by

$$c=c_0-\beta s+\delta s^3, \qquad \phi=c_0 s-\tfrac{1}{2}\beta s^2+\tfrac{1}{4}\delta s^4. \tag{3}$$

In order to determine the proper value of β we can visualize the effect of focusing as follows: in the right-hand half of the curve the cubic error shifts all grooves to the right, the linear error shifts them to the left, so that near the center a groove is to the left of its true position, while farther out it is still shifted to the right. A very close approximation to the best focus will be when the greatest shifts to the right and left are equal, for then the absolute value of the shifts will be a minimum. The phase is a maximum when $c=0$, or $s^2=\frac{\beta}{\delta}$. Substituting this value of s in (2), we have

$$\phi_{\max}=-\tfrac{1}{4}\frac{\beta^2}{\delta}\theta.$$

At the end of the curve

$$\phi=-\tfrac{1}{2}\beta s_0^2+\tfrac{1}{4}\delta s_0^4;$$

equating this to $-\phi_{\max}$, we get a quadratic from which we derive

$$\beta=\delta(1/2-1)s_0^2=0.414\delta s_0^2$$

and for this value

$$\phi_{max}=\frac{3-2\sqrt{2}}{4}\delta s_0^4=0.172\frac{\delta s_0^4}{4}.$$

Taking this value therefore is equivalent to reducing the greatest displacement of a groove to 0.172 of its former value. The dotted curves in Fig. 2 show the corresponding modifications of the vector spirals for $c_0=0$ (only half of each spiral is shown). The corresponding diffraction patterns are shown in the dotted curves in Fig. 3. The compensation in (*a*) is almost perfect, the maximum intensity being 0.984; in (*b*) it is still good, the maximum being 0.933; while in (*c*) it is 0.796, or somewhat better than for (*a*) without focusing. We thus see that even when the groove shift is as much as a whole interval, the grating will behave tolerably well when focused.

Odd errors of higher degree.—From the foregoing it is easy to see how we should expect the terms of the 5th, 7th, . . . degree to affect the definition. For the same value of the maximum shift h, the effect will diminish as the degree increases; on the other hand, the compensation by focusing is less perfect. We can form some estimate of the extent of this compensation in a manner similar to that used above for the cubic error. We thus find that for an error of the fifth and seventh degree, h can be reduced respectively to 0.25 and 0.297 of its original value.

Errors of even degree.—The effect of errors which are symmetrical with respect to the center of the grating is quite different from those we have just been considering. In the first place, the vector curve is symmetrical with respect to the normal at its middle point; the curve thus closes completely for some value of c_0, and the minima are in general "clean." (The only exceptions to this arise when, for large errors, two minima and a maximum coalesce to form a single minimum, which may then have some illumination.) In the second place the diffraction pattern is unsymmetrical, because as c_0 increases or decreases, the curve on the one hand is curling up, and on the other is straightening out. It follows from this that in the third place the maximum illumination is not at $c_0=0$, but at some value of c_0 opposite in sign to the curvature produced by the

error. To fix ideas, let the error be positive, that is, let the grating interval increase toward the ends of the grating. Then the maximum occurs for some negative value of c_0, that is, nearer the central image. It might be inferred that the maxima are nearer the central image than their true position, but this is not the case, for the mean spacing of the lines is not the same as the spacing at the center. The true position of the maximum is given, as we have said, by that vector curve the tangents at whose ends are parallel. An inspection of the curves will readily show that this curve corresponds to a larger negative value of c_0 than that for the maximum, so that the displacement of the maximum is really away from the central image for a positive error and hence toward it for a negative error.

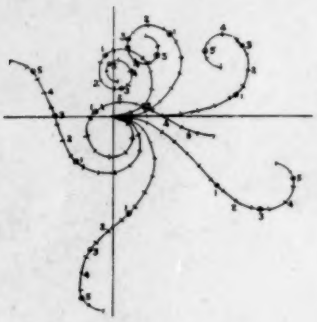

FIG. 4

The most important error of this class, and the only one that will be considered in detail, is the quadratic error. The vector spiral is given by

$$c = c_0 + \gamma s^2$$

or

$$\phi = c_0 s + \tfrac{1}{3}\gamma s^3.$$

Some of these curves are shown in Fig. 4. The full length 15 of this spiral corresponds to

$$\tfrac{1}{3}\gamma s_0^3 = 2\pi.$$

That is, compared with a grating of spacing a, the end grooves are displaced a whole interval. Compared with a grating of the *same mean spacing*, however, the end grooves are in their true position, and the greatest displacement of a groove from its true position is shown by a simple calculation to be

$$\tfrac{1}{3}\gamma s_0^3 \cdot \frac{2}{9}\sqrt{3} = 0.385 \cdot \tfrac{1}{3}\gamma s_0^3,$$

or about three-eighths of the value found by the first process.

The diffraction patterns corresponding to these spirals are given in Fig. 5. The lengths 7.5, 9, 10.5, 12, 13.5, 15 correspond to the curves 1, 2, 3, 4, 5, 6 respectively; the first five of these lengths are indicated by the small numbers in Fig. 4, the full length of the spiral corresponding to curve 6. Similar curves are given by Rayleigh in a discussion of the aberrations of optical instruments, but as we have drawn them here they are shifted in the opposite sense, owing to the different choice of the zero of phase mentioned above. As all of the foregoing treatment applies equally well to the aberrations of an optical instrument, it is worth while showing from that point of view just how this choice of zero differs. A wave-front *ABC* possessing third-order aberrations has different curvatures in the two halves *AB* and *BC* (*B* being the apex of the wave). Rayleigh's choice amounts to taking the normal to the wave-front at *B* as the axis of the instrument, whereas we would take the line from *B* to a point equidistant from *A*, *B*, and *C*.

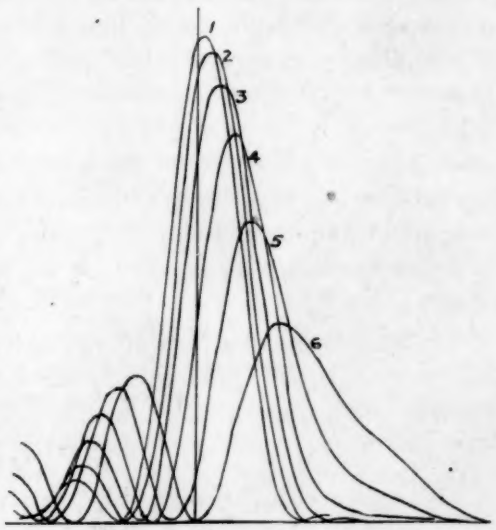

FIG. 5

The series of curves in Fig. 5 show how the quadratic error affects the accuracy of the grating in the absolute measurement of wave-lengths, and shows, if demonstration were necessary, the unreliability of the instrument for this purpose. The displacement of the maximum is very closely proportional to the error of spacing, and amounts to about $\frac{1}{6}\gamma s^3$. Thus the largest error considered here would introduce errors of about $1/R$, where R is the resolving power of the instrument. The error introduced in the method of coincidences comes from another source, and is, as we shall see, of a much higher order of small quantities.

Just as in an optical instrument the aberrations of the third order are those that give most trouble, so in gratings it is these errors of even degree in spacing that constitute the greatest defect. If, therefore, there were some convenient way of introducing such errors artificially, we could rid ourselves completely of the terms of the second degree, and to a considerable extent of those of higher degree. Rayleigh has suggested a slight rotation in azimuth of the telescope or collimator lens. This might prove practicable, but a possible source of trouble arises due to the dispersion of the lens, making it impossible to compensate simultaneously for two coincident spectra.

The presence of the quadratic error may be detected by examining the two halves of the grating separately. The two halves will have different foci, and also the spectra will be less displaced. If in the expression

$$c = c_0 + \gamma s^2$$

we consider each half separately, we have a quadratic error of one-fourth the original amount, combined with equal and opposite linear errors for which

$$\beta = \tfrac{1}{2}\gamma.$$

If we assume with Rayleigh that the change in focus is just perceptible when it involves a change of phase of a half-period, we would just be able to detect this change for

$$\gamma s_0^3 = 4\pi,$$

an error represented in our diagram approximately by curve 4, Fig. 5. The shift of the maximum would be reduced to one-fourth its original value, measured in terms of the width of the spectral image, but as the new images would be twice as wide, the shift in terms of angular measure would only be cut in half. Taking Wadsworth's figure for the possible accuracy of measurement, we could just detect this shift when the error amounts to one-sixth of a grating interval. Apparently, then, the observation of the shift should be more sensitive than that of the change of focus. Taking therefore

$$\tfrac{1}{3}\gamma s_0^3 = \frac{\pi}{2}$$

as the smallest value we could detect, and remembering that the greatest shift of a groove from its true position is only three-eighths of this, we see that it should be possible to detect a quadratic error involving a maximum groove-shift of $0.1a$ by observations on the spectrum of the first order.

Discontinuous errors.—Michelson has called attention to the possibility of a sudden "jump" in the ruling such that the spacing on each side is uniform, but that one interval has the value $a+qa$, where q may be positive or negative. The possible cause of such a discontinuity is given as the release of strains in the speculum metal by the ruling. One may be permitted some skepticism as to the fact of this kind of error having actually occurred. Speculum plates which are figured and polished in one operation do show evidence of strain by a gradual change of figure in the course of a year or so, so that it is necessary to season the plates before they are given their final figure if permanence is desired. The release of strain is gradual, however, and it is difficult to believe that such a change would take place in the ruling without a corresponding change in the figure of the surface which would render the grating unusable. Other possible causes of such a jump in the ruling might be adduced, however, so that it is well to bear it in mind. The theory of this type of error is given by Michelson, as well as the expression for the intensity when the jump is in the middle of the ruling. These results are obtained so readily by the vector method that it is worth while to give the theory from this point of view.

On each side of the break the spacing is uniform; the corresponding portions of the vector curve are therefore circular arcs of the same radius. For the spectrum of the kth order the circular arcs make the constant angle

$$2\pi kq = 2\theta$$

with one another. Fig. 6 shows a set of the vector curves for the case in which the discontinuity is in the middle of the ruling. For all curves of the family the centers of the arcs lie on the lines OP, OQ making an angle of $2\pi kq$ with one another. The phase is the

angle OPB. If ON is the bisector of POQ, the perpendicular BH on ON represents half the amplitude. We have

$$\text{angle } BON = \tfrac{1}{2}\left(\frac{\pi}{2} - \phi - \theta\right)$$

$$BH = OB \cos \tfrac{1}{2}(\phi+\theta) = 2OP \sin \frac{\phi}{2} \cos \tfrac{1}{2}(\phi+\theta) = s_0 \frac{\sin \frac{\phi}{2}}{\frac{\phi}{2}} \cos \tfrac{1}{2}(\phi+\theta),$$

agreeing with Michelson's result except for a change in notation. It is easy to calculate, by simple trigonometry, the resulting

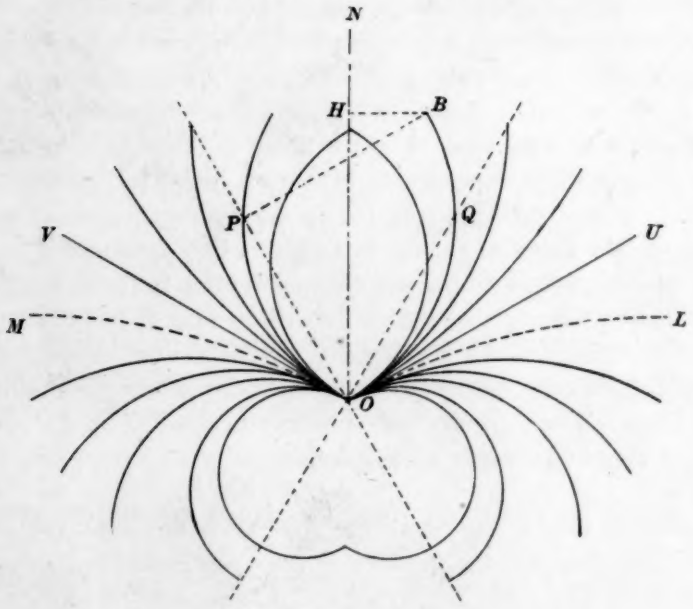

FIG. 6

intensity for any number of breaks at any points of the ruling by this method; the use of graphic methods is probably a simpler way of obtaining numerical results.

There are two sets of minima. The more important of these, and the only one given by Michelson, corresponds to the zeros of the factor $\cos \frac{1}{2}(\phi+\theta)$. These occur when the halves of the vector

curve meet on the y-axis. The other set is given by the zeros of $\sin \frac{\phi}{2}$. These occur when each half of the vector curve closes independently. The first minimum obviously belongs to the first set. The values of ϕ corresponding to the maxima are given by Michelson as approximately $3(\theta - 2\pi n)/4$. These values, however, are with respect to the true direction for a grating with interval a; that is, the zero of phase is taken to correspond to the vector curve consisting of the lines OU, OV. The proper zero of phase corresponds to the dotted arcs OL, OM, so that the maxima are given by subtracting θ from the above. The diffraction patterns are given by Michelson. They present all the characteristic peculiarities of the errors of even degree, though the error cannot be represented by an algebraic expansion, being of infinite degree. If the break occurs at some point other than the middle of the ruling, the error is partly odd and partly even, so that the asymmetry is diminished.

Periodic errors.—The theory of the "ghosts" produced by periodic errors in the spacing has been so often investigated that little can be added to the quantitative aspects of the theory. Michelson's method in particular, employing Fourier expansions for the law of spacing, is peculiarly adapted to the treatment of this kind of error. The vector method, however, gives us immediate insight into the qualitative aspects of the problem, enabling us at once to deduce all the important properties of this class of errors.

Let the $2N$ grooves of the grating consist of a group of m grooves repeated p times. Consider the diffraction pattern of a single group. The error of spacing of this group may be any kind whatever, provided it is not periodic.

To fix ideas, let the error be of even degree. We shall then have some such series of vector curves as is shown in Fig. 7. The curves, the tangents at whose ends are oppositely directed, correspond to the true direction of the first minima for a grating of the same mean spacing; the curves (a) and (b) correspond similarly to the true direction of the second minima, and so for minima of higher order. It is easy to see that these curves cannot close completely, provided the m grooves cannot be divided into two or more smaller identical groups. If we now take the p groups, we have a grating of the

same mean spacing, and the vector curve for the directions just considered will consist of a repetition of loops similar to the curve for a single group (see Fig. 8). That is, the vector "errors of closure" for the single groups are all in phase. But now the principal maxima are narrowed down to $1/p$ of their former width, and these directions become those of a series of secondary maxima. The amplitude of the ghosts is thus proportional to the amplitude of the error of closure of the single group, and we have seen that this is proportional, within the first order of small quantities, to k, the order of the spectrum. Hence the intensity of the ghosts is proportional to k^2.

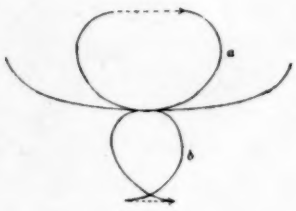

FIG. 7

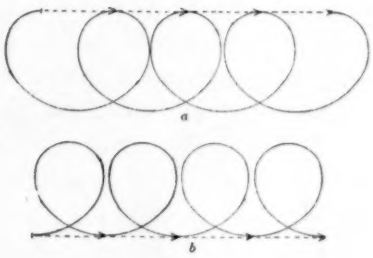

FIG. 8

Accidental errors.—Let us now suppose that the grating interval differs from the mean by a random small quantity $\frac{a\epsilon}{2\pi}$, the values of ϵ being distributed according to the law of error. We can no longer use the continuous vector curve, but must return to the vector polygon. In the direction of the central image the vector polygon is a straight line, in the direction of a principal maximum of the kth order the polygon has become a zigzag line. Consider a direction in which the order of interference is q, where q is in general not integral. The successive sides of the vector polygon have now inclinations $2\pi q+q\epsilon_1$, $4\pi q+q\epsilon_2$, ... We can resolve each elementary vector along and at right angles to its "true direction" (i.e., the direction it would have had if the spacing were perfect). The first part gives us a set of vectors of magnitudes $\cos q\epsilon_1$, $\cos q\epsilon_2$, equally inclined to one another. As $\cos q\epsilon$ differs from unity by small quantities of the second order, the resultant of this group of vectors gives us the amplitude and phase which we would have had if the spacing

had been perfect. The second group gives a set of vectors of magnitude sin $q\epsilon_1$, sin $q\epsilon_2$, ..., or $q\epsilon_1$, $q\epsilon_2$, ..., neglecting second-order quantities, also equally inclined to one another. If now q is not integral, the distribution of these vectors in phase becomes, on the average, symmetrical. The problem becomes that of finding the "expectation" of amplitude of $2N$ vectors of random phase with amplitudes distributed according to the "law of error." The theory of this problem has been given by Rayleigh, and we thus find that the "expectation" of amplitude is

$$\sqrt{2N}q\epsilon_0,$$

where

$$\epsilon_0=\sqrt{\Sigma\epsilon^2}$$

is the measure of precision of the ruling. The effect of the accidental errors will thus be to produce scattered light between the spectra of intensity proportional to $2Nq^2\epsilon_0^2$. The intensity of the scattered light is thus proportional to the mean square error of position of the grooves and to the square of the order of the spectrum. The term "order" should here be taken to mean the order of interference; thus for a point midway between the first and second orders we take $q=1.5$. The intensity of a principal maximum varies as N^2, so that the relative intensity of the scattered light varies as $1/N$. The total energy in a spectrum is, however, proportional to N, so that the proportion of energy in the scattered light is independent of N, as we should expect.

Amplitude errors.—We now pass to the consideration of a fundamentally different class of errors. These arise when the amplitude of the disturbance from the nth groove is a function of n, while the phase-difference between successive grooves is constant. For such errors it is convenient to take ϕ as our independent variable and write

$$\frac{ds}{d\phi}\equiv\rho=f(\phi).$$

The vector polygon is equiangular but not equilateral. In the direction of a principal maximum, therefore, it always straightens out completely, so that the form of the diffraction pattern is independent of the order. It follows also that the spectral images are

always symmetrical and in their true position. The question as to whether the minima are clean or not depends, as with the phase errors, on whether $f(n)$ is an even function of n or not. In the classification of errors of this type we may repeat exactly the same list as was given for phase errors, namely, (1) f is a polynomial in ϕ; (2) f is subject to discontinuous variations; (3) f is periodic; (4) f is subject to accidental variations. Errors of the first class will naturally arise from the gradual wear of the ruling point; other types of this defect belong to other forms of spectroscopic apparatus. Errors of the second type may arise from a sudden breakage of the diamond point during the ruling; there are also certain artificially produced modifications of the grating coming under this head. Properly speaking there is no amplitude error of the periodic type to be looked for in the process of ruling the grating, but we shall consider one case in which such a variation of amplitude may be made to serve a useful purpose. The fourth type of error is always present.

If we screen off a portion of the ends of the grooves of one part of the grating the amplitude of the disturbances coming from these grooves is cut down proportionally. This process of artificially producing amplitude errors not only serves to give a concrete picture of any particular problem, but affords a means of practically compensating all except the accidental errors.

Generally speaking, the progressive change in form of groove due to wear of the ruling point is a trivial defect. The calculation of the intensity when $f(\phi)$ is represented by an algebraic expansion is a matter of some difficulty, but if we write

$$ds = ds_0 e^{-a\phi}$$

we have

$$\rho = \rho_0 e^{a\phi},$$

$$s = \frac{\rho_0}{a}(e^{-a\phi} - 1),$$

the vector curve is readily seen to be a logarithmic spiral, and the intensity is given by

$$\rho_0^2 \cdot \frac{e^{-2a\Phi} - 2e^{-a\Phi}\cos\Phi + 1}{1 + a^2},$$

where Φ is the total range of phase. The form of the diffraction pattern is that of a prism with finite absorption, a case already treated by H. M. Reese,[1] Wadsworth,[2] and the writer.[3] It is there found that the change in resolving power for a total variation of amplitude of $1/e$ is almost negligible, and even for a variation of $1/e^2$ is only 4 per cent. While the wear of the ruling point would not follow any such law in general, the above figures serve as a fair index of the order of magnitude of the effect to be expected. The breakage of the point, producing a sudden change in amplitude, is easily treated by the vector diagram. Thus suppose we have half of the grating giving an amplitude s and the other half giving an amplitude μs. The appropriate vector curve consists of two circular arcs of radii $\frac{s}{\phi}$ and $\frac{\mu s}{\phi}$, and of lengths s and μs, having a common tangent at their point of junction. The expression for the intensity is readily seen to be

$$\frac{s^2}{\phi^2}(1-2\mu\cos\phi+\mu^2),$$

or in general, for lengths x_1, x_2 of ruled surface, with amplitudes A_1, A_2,

$$I=A_1^2+A_2^2-2A_1A_2\cos\Phi.$$

The use of screens to cut off a portion of the ruled surface has a practical importance to which Rayleigh has called attention, but which has not been utilized to the fullest extent. We will consider some of the general aspects of the problem, limiting the discussion to symmetrical screens, for which the minima are free from illumination. Suppose first that the amplitude is reduced at the ends of the ruled surface. The circular arcs which constitute the vector curve of the complete grating now have a smaller radius of curvature at their extremities. Thus in the true direction of the first minimum the vector curve has not closed completely, the closure taking place for a larger value of the phase. The central portion of the image is thus widened out, with a consequent loss of resolving power. On the other hand, the linear dimensions of the closed loop

[1] *Astrophysical Journal*, **13**, 199, 1901.

[2] *Phil. Mag.* (6), **5**, 355, 1903.

[3] *Astrophysical Journal*, **44**, 76, 1916.

are smaller than those of the corresponding circle; hence when the first subsidiary maximum has been reached it has a smaller relative amplitude than the complete grating. The general effect of screening the ends of the grating is thus to widen the image and at the same time to diminish the relative prominence of the secondary maxima. Screening the central portions of the image produces exactly the opposite effect. It is unnecessary to repeat the complement of the preceding argument. Diffraction patterns for discontinuous screening of the center are given by Rayleigh.[1] A typical example of the effect of screening the ends arises when the aperture of the optical train is less than the aperture of the grating. This case has in effect been worked out by Airy and by Rayleigh, the intensity being expressed by means of the Bessel function of the first order. Finally, if we screen off portions of the grating intermediate between the middle and the ends, we get a diffraction pattern which is in a sense a compromise between the two extreme cases given above, in which we obtain an increased resolving power without unduly enhancing the prominence of the subsidiary maxima. This case is discussed by Rayleigh,[2] who finds that if at the points of trisection of the grating three-eighths of the ruled surface is blotted out we get a fairly satisfactory compromise. The vector curves are readily constructed for this case, and the analytic expression for the intensity is readily obtained from such diagrams. The problem of finding the best compromise deserves wider investigation, but unfortunately the methods available for its solution seem to be almost entirely empirical. It may be noted in this connection that the grooves most affected by the phase errors of even degree are intermediate between the center and the ends of the ruled surface, so that the use of screens would in general improve the symmetry of the image.

Periodic errors of amplitude.—Consider a grating in which every pth groove is missing, what would be the effect on the diffraction pattern? To solve this problem, consider first the vector diagram of the complete grating. In the direction of the qth minimum the sum of the disturbances from all the grooves is zero. If we draw all the vectors from a common origin we will have $2N$ vectors sym-

[1] *Coll. Papers*, I, 415. [2] *Ibid.*, III, 88.

metrically distributed about this point. If now we obliterate every pth groove, the corresponding vector diagram will be obtained by taking away $2N/p$ vectors, which will be in general symmetrically distributed about the origin, and which will therefore have zero for their sum. The sum of the remaining vectors will thus be zero, so that in general the minima of the defective grating will coincide with those of the complete grating. An exception arises when $q = 2N/p$, $4N/p$, etc., for in this case the vectors subtracted will all be in phase, and will have a resultant $2N/p$; the resultant of the remaining vectors will have the same magnitude. The effect of omitting the grooves will therefore be to produce a series of ghosts of equal brightness $1/p^2$ in the directions of the principal maxima of a grating having an interval pa.

This device has been suggested to the writer by Dr. J. A. Anderson,[1] of Mount Wilson, as a means of obtaining secondary standards for wave-length measurement with a grating, using the method of coincidences. It would make possible the use of a smaller number of selected spectral lines of known homogeneity, for purposes of comparison.

Accidental errors of amplitude.—These are due largely to the crystalline structure of the speculum combined with its brittleness. The range of the variation of amplitude is usually greater for the smaller wave-lengths. Let the amplitude of the successive disturbances differ from a certain mean value A by random quantities ϵA. Consider any direction in which the illumination would be zero for a perfect grating. For the actual grating the vectors are distributed symmetrically in phase, but are of unequal magnitude. We can break up this sum into a sum of $2N$ vectors of amplitude A plus a sum of $2N$ vectors of random amplitudes distributed symmetrically in phase. The sum of the first set is zero, that of the second, by the results of Rayleigh, referred to above, is on the average $A\sqrt{\Sigma\epsilon^2}$. Hence the expectation of intensity is $A^2\epsilon_0^2 N$. The effect of the accidental errors of amplitude is to produce scattered light of an intensity independent of the order. Practically all

[1] The writer is indebted to Dr. J. A. Anderson for his permission to refer to this. Dr. Anderson has actually ruled a small grating of this type, but the results obtained have not yet been published.

gratings show this scattered light, which can be distinguished from that due to errors of spacing by examining the neighborhood of the central image, where the errors of spacing produce no effect. It seems to be perfectly feasible to determine, by measuring the intensity of this scattered light, the magnitude of the accidental errors, both of spacing and groove form.

The method of coincidences.—The discovery of inadmissible errors in Rowland's table of wave-lengths threw a great deal of suspicion on the method of coincidences as a means of constructing standard tables. Michelson[1] then showed that a particular kind of error of spacing would be sufficient to account for the errors actually found. Unless it can be shown, however, that such defects as are unavoidable in the best gratings will produce systematic errors of measurement, a complete case can hardly be said to be made out against the method. Miss Howell's[2] very careful work has reopened the question (1) by pointing out other sources of error which were present in Rowland's work; (2) by showing that with some gratings, at least, good measurements can be made. Our task here is to investigate just what defects impair the usefulness of gratings for accurate measurement, together with the size of the possible errors introduced.

To begin with we may notice that coincident spectra of different orders will be in focus simultaneously. Consider the wave-fronts of two coincident orders. The effective errors of spacing are proportional to the order, and hence vary inversely as the wave-lengths, so that the phase variations of the wave-front are in this ratio. The linear magnitude of the distortions of the wave-front are thus the same, or the two wave-fronts coincide.

In the second place, variations of the form of groove or other amplitude errors cannot affect the measurements, as they affect neither the symmetry nor the positions of the spectral images. The same is true of the phase errors of odd degree. There remain the phase errors of even degree as the only defects which can affect the measurements with any systematic error.

We have seen that such errors displace the maxima from their true positions. We shall limit the following discussion to quad-

[1] *Astrophysical Journal*, **18**, 298, 1903. [2] *Ibid.*, **29**, 230, 1914.

ratic errors. If we plot the position of the maximum against the magnitude of $\frac{1}{3}\gamma s_0^3$, the coefficient of the quadratic error, we get the curve (*a*) of Fig. 9. Except for the larger values of $\frac{1}{3}\gamma s_0^3$ this curve is sensibly a straight line. Thus the displacement of the maximum, measured in phase (that is, in terms of the width of the image),

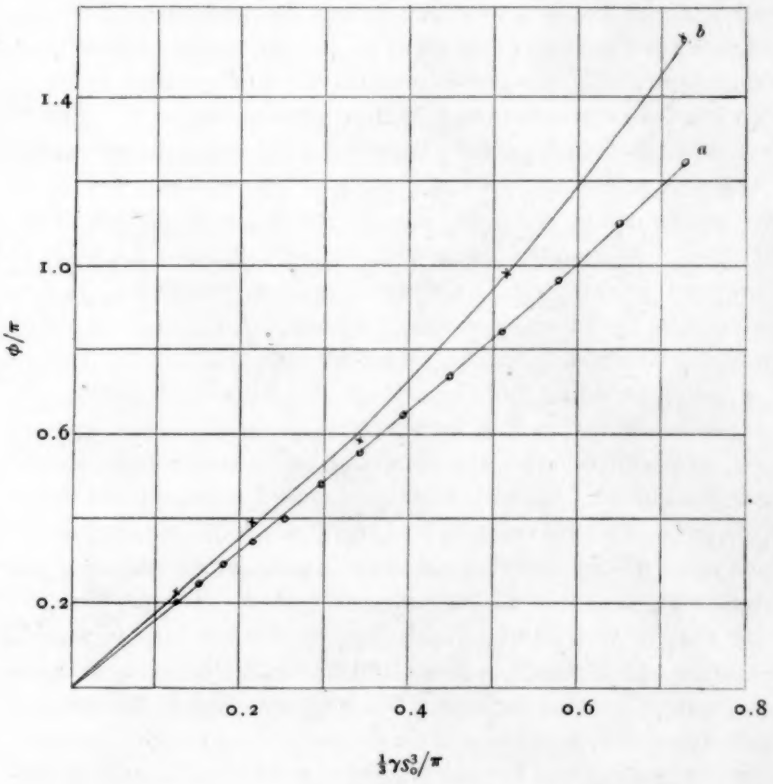

FIG. 9

is proportional to the magnitude of the error. The error is proportional to the order of the spectrum, while the width of the image is proportional to the wave-length; the angular displacement of the maximum is therefore proportional to $k\lambda$. It follows that if all measurements are made on the maxima, no errors can be introduced in the ratio of two wave-lengths. The same is true if we

make all settings on the center of gravity of the central maximum, for if we plot the shift of this we get the curve (*b*), which is approximately straight. The error committed arises entirely from the fact that the images are unsymmetrical, and dissimilar for the different orders, and hence when we set on an image, the particular portion of the image set on depends on the form of the image, and on the intensity of the line in a way that cannot be predicted. The distance between the two curves gives us an approximate upper limit to the magnitude of the error committed. The greatest value of this within the range of errors considered gives about $\pi/6$. This is about one-sixth of the angular separation of the principal maximum and the first minimum. For a grating of 100,000 lines the error would be about 2 in 1,000,000, say 0.01 A in the neighborhood of the D lines. This would suppose, however, a grating so bad that the spectral images were nearly symmetrical doublets. If the asymmetry of the images is no greater than 1 of Fig. 5 the limit of error would be about 0.001 A. Whether such gratings are hard to find or not an asymmetry of this amount should be easily noticeable, so that it should be possible, by simple examination of the spectral images, to decide whether the results of the measurements should have this accuracy. There are other tests for the presence of quadratic error, such as covering half of the grating and measuring the change of focus or observing the shift of the image, but they are not nearly so sensitive as the asymmetry test. It may be considered that in view of the multiplication of secondary standards obtained by the Fabry and Perot interferometer there is no need to use gratings for this purpose. We shall see in what follows that the interferometer measurements are subject to an error of the same kind as the grating, and further, the grating remains the only instrument for measurements on absorption lines. The writer hopes at some future date, when he can get access to a number of gratings, to pursue the subject farther.

The defective Fabry and Perot interferometer.—Consider a narrow beam of parallel light falling on the interferometer plates. Let Δs_0 be the amplitude of the first transmitted beam. Then the amplitude of the nth beam may be written

$$\Delta s = \Delta s_0 e^{-kn}.$$

The vector polygon consists of a series of vectors whose phases are in arithmetical progression and whose amplitudes are in geometric progression. The summation leads to the well-known Airy formula for the intensity. If the logarithmic decrement of the successive interfering beams is not too great, the writer has shown that we can replace the polygon by a logarithmic spiral. Hence we may write

$$ds = ds_0 e^{-kn} = \rho_0 e^{-\frac{k\phi}{d\phi}} d\phi = \rho_0 e^{-\frac{k\phi}{c_0 ds_0}} d\phi .$$

Hence

$$\int_0^\infty \cos\phi ds = \frac{k}{c_0^2 ds_0^2 + k^2} \qquad \int_0^\infty \sin\phi ds = \frac{c_0 ds_0}{c_0^2 ds_0^2 + k^2} .$$

Here c_0, the curvature of the spiral, is the change in phase for unit length of the spiral. Hence $c_0 ds_0$ is the phase-difference between successive interfering beams ($=\phi_1$, say), while k is the logarithmic decrement for two successive beams. Squaring and adding, we get the expression for the intensity previously obtained.[1] If the beam covers an area dS of the plates, it contributes to the total intensity the two vector components

$$\frac{kdS}{\phi_1^2 + k^2}, \quad \frac{\phi_1 dS}{\phi_1^2 + k^2} .$$

Let the distance z between the plates be given by

$$z = z_0 + f(x, y),$$

where z is the distance at any assumed origin on the plates, and x, y are the co-ordinates of dS. We have

$$\frac{2\pi}{\lambda} z = \phi_1 = \Phi + \frac{2\pi f}{\lambda} ,$$

where Φ is the phase-difference at the origin. We thus get for the final intensity the sum of the squares of the two integrals

$$\int \frac{kdS}{\left(\Phi + \frac{2\pi f}{\lambda}\right)^2 + k^2}, \qquad \int \frac{\left(\Phi + \frac{2\pi f}{\lambda}\right) dS}{\left(\Phi + \frac{2\pi f}{\lambda}\right)^2 + k^2} .$$

[1] Sparrow, *op. cit.*

The images will not in general be symmetrical. Before examining the general case we will assume a few special forms of f. Taking circular plates, with the origin at the center, let

$$z=z_0+\frac{\lambda b r^2}{2\pi R^2},$$

where R is the radius of the plates, we can readily integrate, and get for the intensity

$$J=\frac{k^2}{b^2}\left\{\tan^{-1}\frac{bk}{k^2+\phi_1^2-\frac{b^2}{4}}\right\}^2+\frac{k^2}{2b^2}\left\{\log\frac{\left(\phi_1+\frac{b}{2}\right)^2+k^2}{\left(\phi_1-\frac{b}{2}\right)^2+k^2}\right\}^2$$

The images are symmetrical. The curve is shown in (b) Fig. 10. Perhaps the simplest assumption we can make as to f is that f is constant and equal to 0 for the fraction h of the total area, and equal to the constant $\gamma\frac{\lambda}{2\pi}$ for the remainder. The expression for the intensity is

$$J=\left[\frac{hk}{\phi_1^2+k^2}+\frac{(1-h)k}{(\phi_1+\gamma)^2+k^2}\right]^2+\left[\frac{h\phi_1}{\phi_1^2+k^2}+\frac{(1-h)(\phi_1+\gamma)}{(\phi_1+\gamma)^2+k^2}\right]^2$$

For $h=\frac{1}{2}$ we get a symmetrical curve. For $h=\frac{1}{3}$ we get the curves (c), (d), (e), for which $\gamma=0.18$, 0.27, 0.36 respectively.

Before examining the consequences of this asymmetry, let us look a moment at the general question. In the first place, it is evident that the number of surfaces which will give the same diffraction pattern is infinite, for we can rearrange the elements dS in any way so long as the distances are preserved. If we change the sign of ϕ_1 and f simultaneously, the pattern is unaltered. But changing the sign of f replaces the surface by its reflection in the xy plane, and since this is equivalent to a change in the sign of ϕ_1, the diffraction pattern is turned end for end. If the pattern is symmetrical therefore, f and $-f$ give the same pattern. Thus a concave and a convex spherical curvature give the same pattern. Choose z as independent variable, and pass planes z, $z+dz$, intercepting between them the area dS. Let $\bar{z}$ be the plane through the

center of gravity of the surface. Then if dS/dz is an even function of $z-\bar{z}$, the elements are *balanced* against one another, so that when the surface is reflected in the plane $z=\bar{z}$, there results simply an interchange of equal elements, and we have symmetry. It is not quite clear whether this condition is necessary for symmetry, as the difficulties of a general solution of the problem are considerable. We can see now, however, why the spherical curvature should give symmetrical images for this condition is exactly fulfilled in this case, dS/dz being in fact constant. With the notation adopted, the quantity b represents the extreme deviation in phase over the plates.

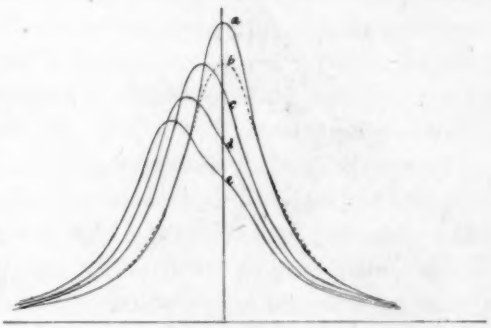

FIG. 10

By examining the expression for the intensity it can be seen that the shape of the diffraction curve depends on the ratio of γ to k. The value of k corresponding to the investigations of Fabry and Perot is log 0.75 or about 0.3, which is therefore the value assumed in these calculations. The better methods of polishing the films used by Pfund and others probably reduce this value of k, and hence make higher demands on the accuracy of the plates. Thus the errors of path difference corresponding to the curves given are (*a*) 0; (*b*) 0.05λ; (*c*) 0.032λ; (*d*) 0.043λ; (*e*) 0.064λ.

The importance of this for us lies in the fact that the method of measuring wave-lengths with this instrument is essentially the *method of coincidences*. The ratio of two wave-lengths is found by measuring the separations of the plates in terms of the two. If the images are unsymmetrical, an error arises, just as with the grating, from the uncertainty as to which part of the fringe is set on. The distance between the center of gravity and the maximum for the curves drawn is (*a*) 0; (*b*) 0; (*c*) 0.05; (*d*) 0.10; (*e*) 0.14.

The separation of the fringes is 2π in terms of ϕ. Thus if the variation of path difference is 0.04λ the error of setting may amount

to one-sixtieth the distance between the fringes, or about double the limit of measurable accuracy. But this requires that the distance between the plates be constant to within $\lambda/40$, or that each plate be correct to within half this amount if the plates are tested separately. We thus see what a high precision of figure is necessary to approach the accuracy which the instrument is supposed to give. Even if the plates are figured accurately a slight variation in the condition or thickness of the silver film may introduce the errors. Somewhat larger errors are caused when $h=\frac{1}{4}$. This is probably an extreme case, but the value $\frac{1}{3}$ does not give the upper limit of possible asymmetry.

It may be argued, however, that the method of using different separations of the plates, which is used in order to eliminate the phase-change in the film, also eliminates any systematic error in setting on the fringes. In the first place, in order that this method should be valid, the two plates when used with different etalons must be always placed in the same relative position. In the second place, the wave-length must be calculated from the *difference* of the separations, just as is done in absolute measurements. Actually, two different methods have been used. Pfund[1] calculates the phase-change from the difference, which thus gives the average value of the phase-change over the plates, and then uses the larger separation alone to calculate the wave-length, thus reintroducing any error arising from asymmetry. Fabry and Buisson[2] used for the smaller separation the Newton ring system localized *in the plates*. In this system of fringes each fringe is produced by that small portion of the plates in its neighborhood, and the fringes are necessarily symmetrical. Thus in determining the phase change by this method, measurements on symmetrical fringes are compared with those on (possibly) asymmetrical fringes.

There is another cause of asymmetry which exists even with perfect plates. This is due to the fact that the diameters of successive rings vary as the square roots of the natural numbers. The dispersion of the instrument thus varies, and the asymmetry of the rings increases as we approach the center. Thus if we measure the central rings, as is usually done, we have to deal with an asymmetry

[1] *Astrophysical Journal*, **28**, 197, 1908. [2] *Ibid.*, 169, 1908.

the amount of which depends on the fractional part of the order. Errors arising from this cause may affect even absolute measurements, unless care is taken always to use such distances that the fractional order is the same for both separations of the plates. All of the foregoing discussion of the shape of the plates has assumed that the rings were not too near the center. To find the exact shape of a particular fringe, we should have to replace ϕ_1 by μ_2 where r is the radius of the fringe corresponding to the phase ϕ_1 and r_0 corresponds to the zero of phase.

We thus see that the errors of the method of coincidences are not confined to the grating, but are to be looked for in the interferometer as well. That the measurements with the latter instrument have been more satisfactory is due perhaps to the fact that good interferometers have been compared with poor gratings, but also, and in large measure, we think, to the fact that the users of the former instrument have applied the knowledge acquired in other fields regarding the standardization of light sources, atmospheric dispersion, etc. The theoretical advantages of the two instruments may be compared as follows: The defects of the gratings are fixed. If we have a good instrument we can make measurements as good as the best; if it is defective a definite limit is set to the accuracy of all our measurements. With the interferometer a good pair of plates may become defective through rotating one of them, or through tarnish or non-uniformity of the silver film; on the other hand, we may more or less average out the errors of a defective pair by this process, since each silvering or change of relative position makes virtually a new pair of plates. We think the tests for the accuracy of the grating can be made more sensitive than those for the interferometer, because the disappearance of the subsidiary maximum on one side in an intensity curve like (1), Fig. 5, is more certainly detected than a slight general asymmetry of a smooth curve like (*c*), Fig. 10. The writer hopes at some future date to examine this question experimentally.

ROUSS PHYSICAL LABORATORY
UNIVERSITY OF VIRGINIA
July 1, 1918

STUDIES BASED ON THE COLORS AND MAGNITUDES IN STELLAR CLUSTERS[1]

TENTH PAPER: A CRITICAL MAGNITUDE IN THE SEQUENCE OF STELLAR LUMINOSITIES

By HARLOW SHAPLEY

The investigations reported in the preceding *Contributions* point to the close equality of the median apparent magnitudes of cluster-type variables as a significant phenomenon in Cepheid variation. The absolute brightness corresponding to this median value marks the limit fainter than which variability of the Cepheid type evidently does not occur. The discussion in the present paper tends to show that the same absolute magnitude, or one nearly the same, also marks a critical point in the dynamical history of stars in general. The result might have been anticipated, for the concentration of the median luminosities of cluster-type variables to a very narrow interval would require a similar concentration for other stars of like mass and color if Cepheid variation is regarded as accidental rather than inevitable in a star's development. We must keep in mind, however, the possibility that the variation itself might operate to hold a star indefinitely long at the critical magnitude.

From an unpublished study of Messier 3 (N.G.C. 5272) the frequency of photo-visual magnitudes for various classes of colors is given in Table I. The catalogue from which this table was constructed is based upon about ten thousand measures of magnitude. Much care has been taken to avoid errors, and the results are believed to be dependable, both as to actual colors and as to relative frequencies. Only stars between 2′ and 11′ from the center of the cluster and brighter than the seventeenth photo-visual magnitude are included in Table I. All known variables are omitted, as well as stars with doubtful colors or magnitudes. The general discussion of the table we reserve for the paper on Messier 3, using for the present investigation only the relation of star numbers to magnitude.

[1] *Contributions from the Mount Wilson Observatory*, No. 155.

TABLE I

FREQUENCY OF PHOTO-VISUAL MAGNITUDES IN MESSIER 3 FOR EACH COLOR-CLASS (DISTANCE FROM CENTER 2'.0 TO 11'.3)*

Limits of Photo-Visual Magnitude	<*b*5	*b*5 to *a*0	*a*0 to *a*5	*a*5 to *f*0	*f*0 to *f*5	*f*5 to *g*0	*g*0 to *g*5	*g*5 to *k*0	*k*0 to *k*5	>*k*5	<*f*5	>*f*5	All Colors
12.20–.39	.....	.....	.....	.....	.....	.....	.....	.....	.....	1	.....	1	1
.40–.59	.....	.....	.....	.....	.....	.....	.....	.....	1	1	.....	2	2
.60–.79	.....	.....	.....	.....	.....	1	.....	1	.....	2	.....	4	4
.80–.99	.....	.....	.....	.....	.....	1	.....	.....	4	.....	.....	5	5
13.00–.19	.....	.....	.....	.....	.....	.....	.....	5	2	.....	.....	7	7
.20–.39	.....	.....	.....	.....	.....	.....	.....	1	2	.....	.....	3	3
.40–.59	.....	.....	.....	.....	.....	1	.....	4	1	.....	.....	6	6
.60–.79	.....	.....	.....	.....	.....	.....	4	3	.....	.....	.....	7	7
.80–.99	.....	.....	.....	.....	1	.....	7	6	.....	.....	1	13	14
14.00–.19	.....	.....	1	.....	.....	2	2	5	1	.....	1	10	11
.20–.39	.....	.....	.....	.....	.....	1	3	3	.....	.....	.....	7	7
.40–.59	.....	1	.....	1	3	16	6	.....	.....	.....	5	22	27
.60–.79	1	.....	.....	1	1	11	11	2	.....	.....	3	24	27
.80–.99	.....	.....	.....	1	1	3	7	4	.....	.....	2	14	16
15.00–.19	.....	.....	1	2	.....	7	17	3	.....	.....	3	27	30
.20–.39	.....	1	4	4	15	19	13	1	.....	.....	24	33	57
.40–.59	.....	9	23	12	15	26	1	.....	.....	.....	59	27	86
.60–.79	.....	8	6	4	2	28	8	2	.....	.....	20	38	58
.80–.99	.....	2	6	.....	7	17	5	.....	.....	.....	15	22	37
16.00–.19	2	2	2	.....	12	14	6	.....	.....	.....	18	20	38
.20–.39	1	1	.....	4	16	13	1	1	.....	.....	22	15	37
.40–.59	.....	1	.....	9	21	19	.....	.....	.....	.....	31	19	50
.60–.79	.....	.....	2	20	26	11	1	.....	.....	.....	48	12	60
.80–.99	.....	2	4	24	38	1	.....	.....	.....	.....	68	1	69
All magnitudes	4	27	49	82	158	191	92	41	11	4	320	339	659

*To save space in the table two tenth-magnitude red stars and one of eleventh magnitude with color-class *a* are omitted. Probably they are not physical members of the cluster.

Plotting the data of the last column, we obtain, as in the upper part of Fig. 1, the luminosity-curve for all colors, which shows a conspicuous, sharp maximum at photo-visual magnitude 15.50. This combines the luminosity-curve of the reddish giants with an

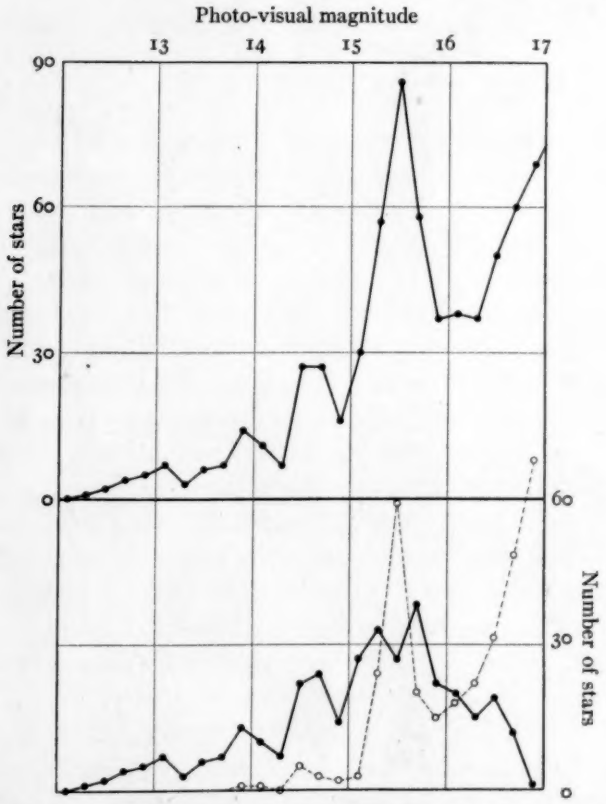

FIG. 1.—Frequency of photo-visual magnitudes in Messier 3. *Above:* All colors. *Below:* Full line, color-index $\geqq +0.60$; broken line, color-index $< +0.60$.

abnormal or at least irregular luminosity-curve of the bluer stars. In the lower half of Fig. 1 the two are given separately. The curve for the redder giants is as smooth and complete as could be expected for a small group of giant stars.[1] The other, though otherwise a regular curve for which the fainter limit is set by the

[1] Actually this curve is also a composite of several distinct luminosity-curves, as may be inferred from the columns of Table I.

limit of the catalogue, shows a remarkable gathering of stars between magnitudes 15.20 and 15.60. A smooth curve would allow less then 15 stars to this interval; the actual number is 83, exclusive of known variable stars.

The median photo-visual magnitude of the variables has not been directly determined, but it is known[1] to lie between 15.25 and 15.35. Therefore, if all the known variables within the measured area were included, with median values of their magnitudes and colors, the excess of stars between 15.20 and 15.60 would be about 1200 per cent.

The observed luminosity-curve for the blue stars is not of the regular form originally expected and most readily explained. If, with the variables excluded, a smoothly increasing curve had resulted, we would have assumed that the variables are peculiar stars intermingled with a normal aggregation. On the other hand, if a smoothly increasing curve resulted only when all the variables were included at their median, maximum, or minimum brightness, we would have assumed that the variables are normal stars undergoing a stage in stellar evolution that is common to nearly all stars with sufficient mass to attain this critical luminosity. The actual condition, however, is accounted for less easily.

A question naturally arises concerning the constancy of the light of the stars that form the excess over the number required for a smooth luminosity-curve. In the area covered by the manuscript catalogue of stars in Messier 3 there are about 100 typical cluster-type variables—all nearly identical in every known characteristic;[2] and there are about 70 of these additional stars, which, though closely resembling the variables in magnitude and color, do not in a single case appear to undergo a definite variation of light. Are they incipient variables, or stars in which a variation once existed but is now disappearing or has completely vanished? Or, through the fortunes of position and motion, have they avoided the cause of variation or, by relatively slight differences in physical structure, have they been able to withstand it?

[1] See Figs. 3 and 4 in the second part of the preceding *Contribution*.

[2] Except, of course, in distance from the center, and possibly, therefore, in space velocity.

The stars in question may be readily selected from the catalogue, and a future detailed study of their magnitudes may help to solve a problem which is of much significance in the full interpretation of Cepheid variation, and which may be of some importance in getting at the internal structure of a star. Meanwhile we have an indication of the possible result through the following tabulation of the means of the photographic residuals:[1]

Photo-visual magnitude 15.20 to 15.60:
Color-index 0.00 to +0.20, Mean Pg. residual ±0.103 (33 stars)
Color-index +0.20 to +0.40, Mean Pg. residual ±0.083 (19 stars)
Color-index +0.40 to +0.60, Mean Pg. residual ±0.071 (37 stars)
All magnitudes and colors, Mean Pg. residual ±0.082 (800 stars)

The interval of brightness near median magnitude is the most comfortable of all for measurement, being neither too bright nor too faint, and the accidental errors should be less than for the average. The mean residual for all stars includes a considerable number of values referring to objects which later were rejected because of uncertainty; therefore this mean is slightly greater than it should be for a fair comparison.

It is evident then that the 33 stars most nearly like the variables in median magnitude and color are subject to uncommonly large magnitude deviations. Moreover, one or two uncertain results are not responsible for the large average residual, as one-half of the deviations exceed ±0.10. While awaiting further study our provisional conclusions are that Cepheid variation affects some of these stars to some extent, and that it may be possible to witness the beginning or end of this type of phenomenon. If so, the initial and final stages must be short compared with the whole duration of mature variability, for of all the variables in this cluster none is yet known to have an intermediate amplitude or period.[2]

Since the mean photographic median magnitude is definitely known from Bailey's extensive studies, we shall be able to examine

[1] The residuals will appear in the forthcoming catalogue of Messier 3. The magnitude of each star is the mean of from two to six determinations.

[2] Cf. sec. IV of the sixth paper. Variable No. 37 belongs to a distinctly different subtype.

more closely the relationship of the variables to these 33 stars if we put the frequency analysis on a photographic basis. Rearranging the data of Table I and plotting the numbers for color-class *a*0–*a*5, we obtain the photographic luminosity-curve given in Fig. *2a*. The maximum frequency corresponds to a brightness fainter than the median, 15.50, by less than a tenth of a magnitude.

A similar investigation of the photographic luminosity-curve for Messier 13—the only other globular cluster for which extensive color

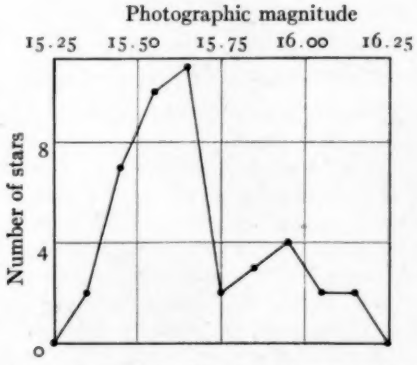

FIG. *2a*.—Photographic luminosity-curve of blue stars in Messier 3.

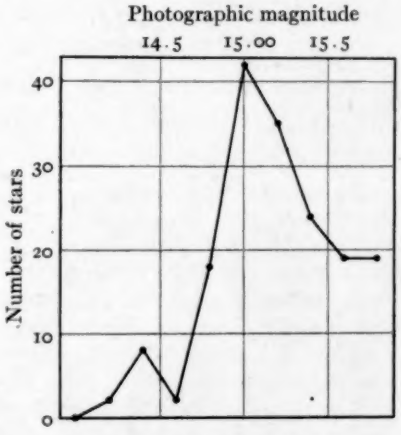

FIG. *2b*.—Photographic luminosity-curve of blue stars in Messier 13.

data are available[1]—results in the curve (Table II) illustrated in Fig. *2b*, showing a similarly close coincidence of maximum frequency and median magnitude (15.03).

In Messier 13 the few faint variable stars have not been thoroughly studied, and it is necessary to obtain a hypothetical median magnitude by adding 1.28 to the mean magnitude of the 25 brightest stars.[2] And also, since the number of stars of color-class *a* is

[1] *Mt. Wilson Contr.*, No. 116, p. 51, 1915. Instead of photographic, as erroneously printed, the limits of magnitude in this table should read photo-visual. The change to the photographic system for the present curve is made with sufficient accuracy by assuming mean colors and magnitudes for each tabulated interval. For Messier 3, however, the transformation was carried out rigorously.

[2] See sec. V of the sixth paper and Table II of the seventh paper.

not large, the interval of color is extended in both directions, thus including numerous stars with negative color-indices. In fact, the maximum is chiefly due to stars with color-class *b*7 to *a*0—a comparatively infrequent type[1] in Messier 3.

TABLE II

FREQUENCY IN MESSIER 13 OF THE PHOTOGRAPHIC MAGNITUDES OF HIGH-LUMINOSITY BLUE STARS

Interval of Photographic Magnitude	*b*0 to *b*5	*b*5 to *a*0	*a*0 to *a*5	*a*5 to *f*0	*f*0 to *f*5	Color-Index < +0.60
12.70–12.89	1	.....	.....	.....	.....	1
.90–13.09	.....	.....	.....	1	.....	1
13.10– .29	.....	.....	.....	.....	1	1
.30– .49	.....	.....	.....	.....	.....	0
.50– .69	.....	.....	.....	.....	.....	0
.70– .89	.....	3	.....	.....	.....	3
.90–14.09	.....	.....	.....	.....	.....	0
14.10– .29	.....	.....	.....	2	.....	2
.30– .49	.....	1	1	.....	6	8
.50– .69	.....	.....	1	.....	1	2
.70– .89	.....	5	6	7	.....	18
.90–15.09	10	24	3	2	3	42
15.10– .29	4	21	9	.....	1	35
.30– .49	1	11	7	3	2	24
.50– .69	.....	5	6	5	3	19
.70– .89	.....	.....	3	4	12	19

The luminosity-curves of Figs. 2*a*, *b* render probable the hypothesis that the median magnitude is an epochal point in the luminosity of a cluster star whether or not its light is periodically variable. As a partial and provisional interpretation of the congestion of blue stars at this magnitude, we may refer to the luminosity-period curve of Cepheid variation and in particular to the constancy of absolute magnitude for all periods less than a day. If we adopt the theory of pulsations for these variables, it is clear that a closely analogous curve would connect luminosity with mean density, and accordingly a progression through the whole range of densities possessed by cluster-type variables would demand no change in median brightness. Now, if a more or less uniform progression of density with time is a property of all stellar evolution (a natural and almost necessary assumption, if stars are either contracting or expanding),

[1] The small difference in the mean color at the maxima of the two curves is probably larger than the combined zero-point errors in the two color scales.

there may come an era in the history of every massive star when no change of magnitude will occur throughout a long interval of time, although the variation of the mean density is making conspicuous progress. Such a condition will result in a maximum in the frequency-curve of luminosity, while it would produce no irregularity in a graph of the frequency of density.[1] Therefore, the perfect counterbalance, at the median magnitude of cluster-type variables, of decreasing surface area with increasing intensity of light-emission, which can account for the flattening out of the luminosity-period curve, may explain at the same time the accumulation of invariable stars about this critical point.[2]

Assuming, however, on the basis of the theoretical studies of Eddington and Jeans, that the brightness of a giant star is independent of the density and is essentially a simple function of the

[1] In the third note of the preceding paper we find in Messier 15 neither a change of color nor an appreciable difference in magnitude for the two groups of variable stars which differ decidedly in period and presumably in mean density. The mean periods are about three and five-eighths of a day, respectively, and the mean densities must differ in the ratio of three to one unless two different classes of vibration are involved. How such a change in the effective density could occur (with its supposedly necessary change in surface area) without a distinct alteration either in the total light-emission or in the median color is not obvious. A detailed study of the two groups of variables in red or violet light may supply the answer.

The solution of this problem may lie in a further consideration of the theoretical discussion and observations of the preceding contribution (*Astrophysical Journal*, **49**, 24, 1919). From equation (2) on page 28 the absolute magnitude may be written

$$M \propto \tfrac{2}{3}(2 \log P + \log \mu) + \log \phi \text{ (C. I.)}$$

where P is the period of pulsation, μ is the mass, and ϕ (C. I.) determines the surface brightness. The observations referred to in the foregoing paragraph suggest that M and ϕ (C. I.) do not change for cluster-type variables. Therefore $P^2\mu$ remains essentially constant, and the variables of the shorter-period subtype are three times as massive, in the mean, as those of the longer-period subtype. Further, within the uncertainty of the observations of color and magnitude, the period of an individual cluster-type variable does not change during the life of the variation, for it is hardly conceivable that the mass alters appreciably. Every period is associated with a definite mass, and only one period is possible for a given star. (Cf. Russell, *Science*, N. S., **49**, 136, 1919.) It appears, therefore, that the period-luminosity curve is best interpreted on the basis of Eddington's theory of a giant star. In any case the critical luminosity is to be associated with the flattening of the period-luminosity curve.

[2] There are large numbers of fainter blue stars in both Messier 3 and Messier 13, as is shown by Table I of this paper and by Tables XXI, XXIV, and XXV of the second paper, *Mt. Wilson Contr.*, No. 116. One is reminded of the differences in mean absolute magnitude found by Kapteyn and Charlier for various groups of galactic B-type stars.

mass, then the congestion of stars at the median magnitude would indicate a sharp maximum in the frequency of stellar masses. Such an assumption suggests interesting and unexpected consequences in the interpretation of the luminosity-period curve[1] and in the study of the relative ages of stellar systems.

Of the two clusters discussed above, one is unusually rich in typical cluster-type variables but poor in *b*-class stars of the median magnitude; the other is poor in variables but remarkably rich in the number of stars of the median magnitude with negative color-indices. There may be much significance in this compensatory difference, but our material is still too scanty to warrant further hypothesis. Future work must also decide to what extent the cluster phenomena may be modified in their translation to the general stellar system.

Meanwhile, through investigations of composite luminosity-curves for other clusters, we may look into the generality of the maximum at the median magnitude. Lacking color-indices, we have not the power to determine quantitatively the influence of the giant red stars on the composite curves, but we shall be able to extend the work to fainter magnitudes and to learn something of the relative importance of this maximum frequency near absolute magnitude zero.[2]

A summary of the photographic magnitudes in four clusters appears in Table III. Messier 13, which affords a good example of a cluster with few variables, has been remeasured for this table because the data of Table II are too incomplete for the fainter stars. For none of the four systems do the measures extend over the whole photographic plate, the condensed central regions, for instance, being necessarily avoided. Over large representative areas, however, all stars are included so that correct relative frequencies are obtained. Care has been taken that peculiarities which might be associated with galactic planes should not unduly prejudice the results. After correcting for distance from the center and irregularities of the measuring scale, the measures are retained in the

[1] See the first note of the preceding *Contribution*.

[2] The adopted absolute value of the photographic median magnitude is −0.23; cf. sec. IV of the sixth paper.

original scale readings, the exact evaluation for fainter stars being impossible from the present material. Except for a possible divergence near the ends of the scale, the average value of an interval is one-third of a magnitude.

The relative luminosity-curves are plotted in Fig. 3, each adjusted so that the median magnitude is on the same vertical line. For both Messier 2 and 13 the median photographic brightness is the hypothetical value based on the brighter stars. The

TABLE III

FREQUENCY OF THE MAGNITUDES OF 4015 STARS IN FOUR CLUSTERS

Corrected Scale Reading	Messier 2	Messier 5	Messier 13	Messier 15
<10	2	1	1	17
10.5	0	0	2	12
11.5	1	0	6	17
12.5	3	2	13	39
13.5	6	10	21	62
14.5	19	26	41	100
15.5	46	45	51	151
16.5	76	79	59	148
17.5	95	109	53	161
18.5	94	126	62	148
19.5	84	110	69	161
20.5	79	89	97	186
21.5	92	79	147	212
22.5	122	90	228	266
Total	719	766	850	1680
Median Pg. Scale	17.2	17.7	14.1	15.4
Median Pg. Mag	15.90	15.26	15.03	15.63
Number of variables included	5	70	3	50

sharpness of all the maxima has probably been somewhat lessened through the process of smoothing out scale irregularities.

In comparing these maxima with the median magnitude it should be remembered that the latter is subject to some uncertainty of measurement and the former to a displacement of unknown amount and direction through the inclusion of the giant red stars. The possible effect of the superposition of the luminosity-curves of the red giants is illustrated in the curve of Fig. 4, where, for Messier 3, the total numbers of Table I are plotted, after transforming them to the photographic system. This curve should be compared with that of Fig. 2*a*.

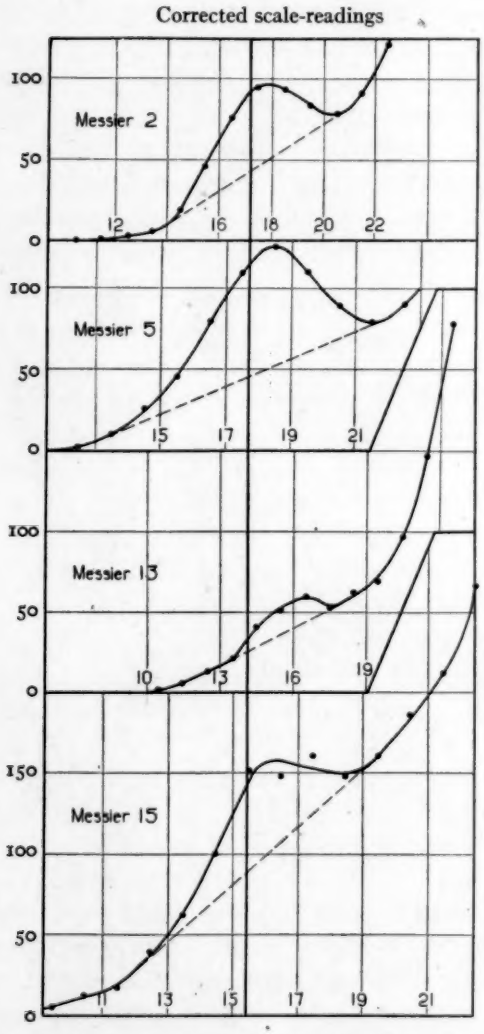

FIG. 3.—Photographic luminosity-curves of stars of all colors in four globular clusters, illustrating the maximum frequency near zero absolute magnitude. The median magnitude for each cluster is indicated by the heavy vertical line. The number of stars measured in different clusters (ordinates) is not proportional to the total number in the cluster.

From the evidence of this series of diagrams there can be little doubt that in all these clusters we have a concentration of the magnitudes of high-luminosity blue stars about the median magnitude of cluster-type variables.

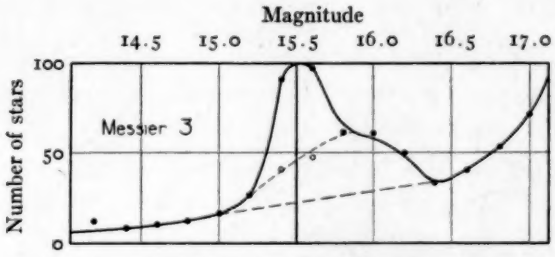

FIG. 4.—Photographic luminosity-curve of stars of all colors in Messier 3. Excluding variable stars we have the broken line with circles.

SUMMARY

The absolute photographic magnitude −0.2 defines a limit of stellar luminosity at which the liability to Cepheid variation suddenly stops. All known variables of the class appear to surpass this limiting brightness, which is probably coincident with a turning-point in the development of the internal structure of a star. Whether or not analogous dynamical changes occur, regardless of variability, at the same epoch in the evolution of luminosity for all giant stars, has been discussed for several globular clusters with the aid of luminosity-curves derived from the magnitudes of 5000 stars. For a graphical representation of the results reference may be made to the figures on the foregoing pages.

Table I affords the complete luminosity-curves for all giant stars in Messier 3 redder than color-class *g*0. Confirming earlier work, the final results for this cluster show that the highly luminous stars are red, that the fainter ones are blue. The infrequence of cluster-type variables in the Hercules cluster (Messier 13) is compensated for by the richness of extremely blue stars at the median magnitude.

MOUNT WILSON SOLAR OBSERVATORY
December 1917

AERIAL PHOTOMETRY

By M. LUCKIESH

Owing to the nature of the material presented in this paper it is impractical to provide the usual summary; therefore for the convenience of the reader the following outline is given of the chief details in their order of appearance in the text.

INTRODUCTION

- Source of daylight illumination
- Variation of sunlight on a clear day
- Variation of skylight on a clear day
- Relations of sunlight and skylight for a clear day
- Relations of sunlight and skylight on four typical days
- Maximum total illumination at noon on clear days as function of season and of cloudiness
- Skylight illumination in terms of total illumination
- Russell's conclusion of 135,000 meter-candles as maximum illumination outside earth's atmosphere
- Another method of determining maximum solar illumination which yields 126,000 meter-candles
- Conclusions

APPARENT REFLECTION FACTORS OF EARTH AREAS

- Terminology—reflection factor and relative brightness
- Table I—reflection factors of fields, barren land, woods, inland water
- General summary of reflection factors of earth areas in summer; in mid-autumn
- Statements of Abbot on earth's albedo

LANDSCAPE

- General characteristics
- Attempt at averaging reflection factors
- Explanation of low reflection factor
- Range of measurements of reflection factor
- Attempt to measure average relative brightness of landscapes

CLOUDS

- Reflection factors of upper surface of different types of clouds
- Explanation of high value

HAZE

Two types of haze
Rough determination of earth haze
Effect of absorption and scattering of haze on earth's relative brightness from high altitudes
Another rough determination of earth haze
False horizon due to earth haze

SKY

General sky brightness
Variation of sky brightness with altitude
Table II—brightness of earth areas in terms of sky brightness
Checks between Tables I and II
Check on mean reflection factor of earth
Determination of relative brightness of upper and lower sides of horizontal opaque white sheet
Case of an opaque white surface disappearing against the sky
Data of Kimball and Thiessen for clear and smoky atmosphere

WATER

Variation of brightness expected from reflection at the surface
Variations due to objects mirrored, to beds, to suspended matter, etc.
Variation of trip over shallow inland, deep inland, and ocean water
Hue considerations
Table III—brightness of inland waters in terms of sky brightness
Same for deep ocean water
Evidence that brightness of water is generally due chiefly to diffusion of light in the water
Numerical considerations
Brightness at forty-five degrees incidence

INTRODUCTION

Some of the data recorded in the following pages were obtained as part of an investigation of the visibility of airplanes. Owing to the necessity for hastening the completion of the work which was primarily in mind, many interesting details could not be given much attention. However, many data were obtained during sixty hours in the air with facilities for making flights of any character as to altitude, distance, terrain, etc., which could not be obtained as satisfactorily in any other manner. Various types of instruments were used, but most of the measurements were made with a compact instrument of special design which need not be described. It is

proposed to present in this paper only certain photometric measurements and allied observations, which may be of interest from various viewpoints aside from the original one. In order that the reader may have a clearer conception of the details described in later sections, a few introductory paragraphs are included which deal with the general characteristics of natural lighting.

Natural light in the daytime reaches the earth directly from the sun and also indirectly through scattering by the atmosphere,

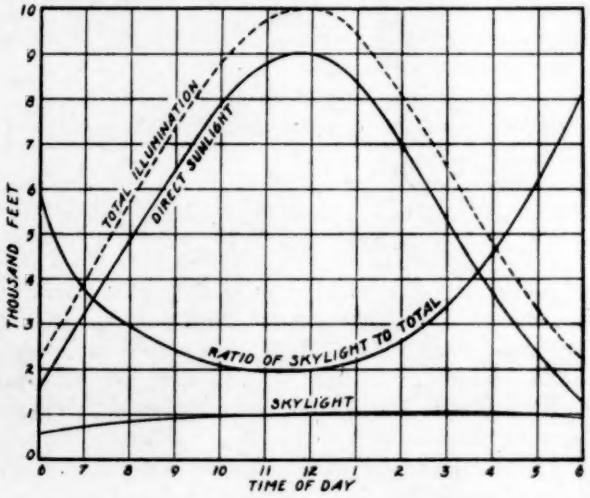

FIG. 1.—Relations of sunlight and skylight on a horizontal surface on a clear day

clouds, haze, etc. For the sake of brevity the former will be termed "sunlight" and the latter "skylight." It will be understood that skylight will mean all the light which reaches a given point from not only the blue sky but also from clouds and haze. If the sky were perfectly clear and if space were a perfect void the illumination on a horizontal surface due to direct sunlight would be proportional to the sine of the sun's altitude and the illumination due to skylight would be zero. Owing to the scattering of light by the atmosphere some skylight is always present. On a clear day the illumination on a horizontal surface due to the sun roughly approximates a sine curve and that due to skylight is approximately constant[1] for a large portion of the day.

[1] M. Luckiesh, *Light and Shade and Their Applications*, 1916, chap. vii.

In Fig. 1 are presented the results obtained at the earth's surface on an exceedingly clear, cloudless day in June when the maximum altitude of the sun was about seventy-four degrees. The measurements represent illumination on a horizontal surface. It is seen that the illumination due to direct sunlight throughout the portion of the day represented approximates a sine curve. The illumination due to skylight is nearly constant, although it increases slightly with the sun's altitude. The ratio of skylight illumination to total illumina-

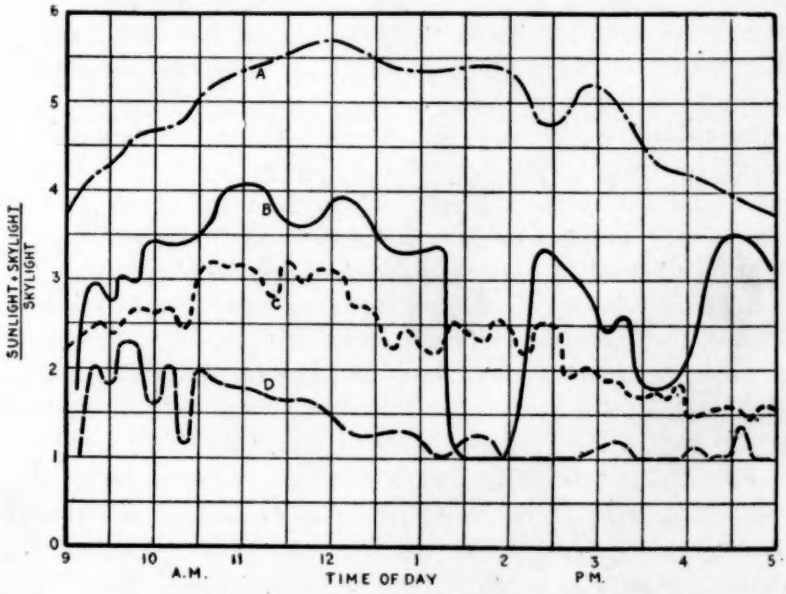

FIG. 2.—Relations of sunlight and skylight on four typical days

tion at noon is 0.1, which represents about the minimum determined at the earth's surface. This ratio generally decreases as the sun's altitude increases. As haziness increases the values of total illumination decrease somewhat and those for skylight illumination increase considerably, with the result that the ratio of skylight to total illumination increases very much.

In Fig. 1 the data represent the results on an ideal, extremely clear and cloudless day. In Fig. 2 the ratios of skylight to total illumination[1] are shown for four common days. Of course, there are many variations owing to clouds, haze, etc. Many such

[1] *Ibid.*

records have been made, but it appears unnecessary to present examples in this paper; however, it is well to visualize the causes for the extreme variability in natural lighting and in the sky aspect.

On perfectly clear days in this country the maximum total illumination on a horizontal surface at the earth at noon varies from about 10,000 foot-candles in June to about 3500 foot-candles in January. According to H. H. Kimball[1] it is less than the direct solar illumination of a normal surface from September to February, inclusive, but exceeds the latter from May to August, inclusive, for a period of from four to eight hours in the middle of the day. On thinly overcast days when the sun is just invisible the total illumination is about one-half that on a clear day with the sun at the same altitude. This value decreases to one-third or one-fourth on thickly overcast days and occasionally reaches a very low value during storms. According to measurements made in various places in this country, on clear days in summer the illumination due to skylight on a horizontal surface varies from one-third to one-tenth of the total illumination on the same surface at noon, and in winter from one-half to one-sixth. A common value or approximate mean for a large part of the year is about one-fifth for the outskirts of Cleveland, Ohio.

H. N. Russell[2] has grouped some of Kimball's data[1] according to the amount of skylight illumination and by extrapolation to the zenith (air mass unity) arrives at a mean transmission factor of the atmosphere equal to 0.77 for 15 clear days. The mean maximum intensity of sunlight from the zenith sun is found by extrapolation to be 9600 foot-candles (103,000 meter-candles) for these same days. The skylight illumination at noon on these days averaged 1000 foot-candles approximately. Assuming the above value of atmospheric transmission factor for visible relation, Russell computes the illumination on a horizontal surface due to zenith sun would be 12,500 foot-candles (134,500 meter-candles) outside our atmosphere.

It is interesting to plot Kimball's value for 22 days, which were grouped by Russell in accordance with the illumination due to skylight alone, that is, according to the haziness of the weather. By plotting values of skylight illumination against those for sunlight

[1] *Monthly Weather Review*, **42**, 648, 1914. [2] *Astrophysical Journal*, **43**, 127, 1916.

illumination it is found by extrapolation that the value for zenith sun illumination on a horizontal surface becomes about 11,700 foot-candles (126,000 meter-candles) when the illumination due to the sky alone reaches zero. This appears to be a method worthy of consideration for determining the maximum possible illumination at the earth that can be expected from the zenith sun.

Notwithstanding the excellent work which has been done, the data pertaining to natural lighting are incomplete. However, it is not the purpose of this paper to discuss the foregoing aspects or to include data pertaining to them. They have been mentioned in order that the reader may appreciate the variables involved in some of the values presented in the sections which follow.

APPARENT REFLECTION FACTORS OF EARTH AREAS

To apply the term reflection factor (commonly called reflection coefficient) to the earth's surface which has "depth" is somewhat misleading; however, as viewed from an appreciable distance above the earth, this term is not misapplied from a practical standpoint. The term "relative brightness" will be used occasionally and, unless otherwise stated, will mean the brightness relative to a horizontal, perfectly reflecting and diffusing opaque white surface under the same illumination. Obviously the reflection factor of this theoretical white surface is equal to unity so the values of "relative brightness" of various earth areas, clouds, sky, etc., are in reality "apparent" reflection factors.

In Table I values of reflection factor are presented for various characteristic earth areas as obtained on various days during late summer. The individual values are means of many observations, and a sufficient number of these are presented to show the consistency of the results. All were made within 150 miles (240 km) of Langley Field at Hampton, Virginia. Further comment on these various types of areas will be found in following sections. Measurements, unless otherwise stated, were made vertically downward from the airplane and usually at altitudes from 1000 to 5000 feet (300 to 1500 meters). The brightness measurements were made with different types of instruments in terms of the brightness of a white diffusely reflecting (or transmitting) surface receiving the

same illumination. These values were then reduced to values in terms of a perfect white. Obviously such measurements had to be made below the clouds when the latter were present. Many measurements were made at various altitudes up to 20,000 feet,

TABLE I

PERCENTAGE OF RELATIVE BRIGHTNESSES OR APPARENT REFLECTION FACTORS OF VARIOUS TYPES OF EARTH AREAS DURING LATE SUMMER

(These were determined in late summer in eastern Virginia within 150 miles of Hampton. In all cases the surfaces were viewed from points directly above them varying from an altitude of 1000 feet upward. The individual values in the following table are means of numerous measurements. In some cases hundreds of measurements were made during a single flight.)

Date		Fields	Barren Land	Woods	Inland Water
Aug. 30, 11:00–12:00	Generally clear	7.4	10.4	4.9	5.7
Aug. 30, 2:00– 3:00	Sunny with large clouds scattered	9.3	10.3	4.9	9.0
Aug. 31, 11:00–12:00	Generally clear, some thin hazy clouds	6.1	10.1	4.4	6.7
Aug. 31, 1:00– 2:00	Fairly clear, some clouds	5.8	11.1	3.8	6.2
Sept. 3, 11:00–12:00	Sunny with large scattered clouds	7.9	15.9	3.9	6.3
Sept. 3, 1:00– 2:00	Generally clear	7.0	19.0	4.6	7.1
Sept. 3, 3:00– 4:30	125-mile trip across country, east and south	6.8	17.1	5.0	6.2
Sept. 4, 11:00–12:00	175-mile trip to Richmond, Petersburg, and return	6.5	13.8	4.5	7.4
Sept. 6, 11:00–12:00	Overcast sky	6.7	13.6	4.4	9.1
Sept. 7, 1:00– 2:00	Clear day	6.3	.....	.....	7.6
Sept. 9, 11:00–12:00	Clear day	.....	11.2	3.6	6.7
Sept. 9, 2:00– 3:00	Clear day	6.7	14.2	3.9	4.4
Sept. 10, 9:00	On trip to Washington and return	5.2	8.7	3.5	5.2
Mean		6.8	13.0	4.3	6.8

Dense clouds, 55 to 78 per cent, for upper sides

Deep ocean water 50 miles from shore, 3.5 per cent

Shadow of cloud on earth's surface is less than 1 per cent of the brightness of the theoretically white surface receiving sunlight and skylight.

although for earth measurements the effect of the low-lying haze was reduced to minimum by making the measurements vertically downward and usually at low altitudes. Most of the measurements were made during the middle of the day because instrumental errors were less at this time. However, no extensive differences were

obtained in the results at various times of the day or even on sunny and overcast days.

Owing to the object of the original work four different types of surfaces were chosen which inclusively represented the entire earth's surface with the exception of cities, towns, and villages, which cover a negligibly small portion. The data in Table I are for summer landscape, and the water areas were entirely inland waters such as Chesapeake Bay, Hampton Roads, and rivers such as the James, the York, the Rappahannock, and the Potomac. Barren land included plowed ground and waste land on which there was no vegetation. Fields included pastures and growing crops.

Most of the observations which are represented by the foregoing mean values on various days were included in the following ranges:

	Percentage
Fields	5–10
Barren land	10–20
Woods	3– 5
Water (inland)	5–10
Deep ocean water	3– 5

A number of flights were made from Langley Field during the middle of October in order to obtain similar measurements for the autumn landscape. The colors were those typical of autumn when the woods were variegated and the fields were dull yellow or brown. These observations were not as exhaustive as those recorded in Table I and on other pages; therefore not much stress is placed on the values. The average values obtained for the apparent reflection factors or the brightnesses relative to a perfectly white surface under the same illumination are as follows:

	Percentage
Fields	6.0
Barren land	11.3
Woods	4.3
Water (inland)	5.5

The value for water in the foregoing does not include many measurements of rivers, as is the case in Table I. In a later section water is discussed further. There is reason to believe that the apparent reflection factors of these earth areas are practically the

same in autumn as in summer. Obviously there should be no difference in the values for water. The landscape during this autumn period was decidedly brownish and generally dull in character. Few green areas were present and the woods displayed a variety of tints. It is noteworthy that the landscape in general appeared more dull and dead from high altitudes than was expected. Although the number of observations during this autumn period was limited to hundreds instead of thousands as in the summer period, it will be noted that the results are not widely different.

It is interesting to note that Abbot[1] from energy measurements concludes that the albedo of a cloud surface is about 0.65, and after allowing for atmospheric absorption computes the albedo of the earth to be 0.60 if it were covered with clouds, 0.14 if it were cloudless, and 0.37 as matters actually are. These values cannot be directly compared with the values of apparent reflection factor given in this paper, however. Abbot's discussion of the reflection factors of the earth is interesting, although the data in this paper will probably modify some of his estimates. It is interesting to compare the values with certain qualifications in mind. For a discussion of albedo two references[2] are recommended.

LANDSCAPE

During the summer period in which the observations whose mean values are presented in Table I were made, the vegetation was predominantly green, although in a few areas there was a slight tinge of yellow common to middle and late summer. There were many plowed fields barren of vegetation and others which were not entirely covered by green foliage. In eastern Virginia there is a great deal of wooded country including wooded swamps, so it appeared, among other reasons, best to keep separate records for the different types of earth areas. Furthermore, there are various kinds of soil, so that the variations in the values in Table I are to be expected. However, the consistency of the observations on woods, water, and fields covered with crops is surprising.

The mean relative brightness or apparent reflection factor for a landscape will depend upon the proportions occupied by the various

[1] *Annals of the Smithsonian Astrophysical Observatory*, **2**, 161.

[2] H. N. Russell, *Astrophysical Journal*, **43**, 103, 173, 1916; Louis Bell, *Ibid.*, **45**, 1, 1917.

types of areas. Attempts were made to estimate the proportions of different types of areas present in the landscapes studied. This could be fairly accurately determined for any given locality by cross-country flights during which a record of time in passing over the various types of areas could be kept. If the airplane speed were approximately constant this would give a fairly satisfactory mean value for the relative brightness or the apparent reflection factor of the terrain. Another plan was to make observations continually and take the average value. This was done a number of times. For example, the mean value obtained during a 175-mile trip from Langley Field to Richmond to Petersburg and return was about 7 per cent. However, it should be noted that this portion of the country contained much wooded area although the course was generally maintained over the more cultivated regions.

Notwithstanding the writer's familiarity with brightness measurements, the generally low values obtained in this investigation were surprising. However, these may be accounted for by considering the earth's surface more in detail. On looking down such areas as fields, woods, etc., one sees a mixture of highlights and shadows. Much of the area receives only skylight, which is a small part of the total light reaching the earth on sunny days. Even on overcast days the shadows receive only a small portion of the total light reaching the earth's surface. Furthermore, the trapping of light is a powerful influence in reducing the reflection factor. For example, a grass area possesses "depth" similar to velvet. Black velvet has a very low reflection factor (about 0.004), but ordinary blotting paper colored with the same dye will possess a diffuse reflection factor of about 2 per cent. The depth of the velvet fiber provides light-traps and shadows, with the result that the reflection factor as a whole is much lower than that of a single fiber. In a similar manner grass plots, corn fields, plowed ground, woods, etc., provide light-traps and shadows so that the mean brightness or apparent reflection factor is materially reduced.

For example, let us assume that the diffuse reflection factor of a blade of grass is 0.16; that the sun contributes 0.8 of the total light; and that the normal aspect of the grass plot is one-half shadow and one-half highlight. If the brightness of the sky is equal to B, then the brightness of a perfectly white horizontal

reflecting surface will equal B due to skylight alone and equal $5\,B$ due to sunlight and skylight. The brightness of the blades of grass receiving both skylight and sunlight will be $5\,B \times 0.16 = 0.8\,B$. The brightness of the grass in shadow, assuming the shadows to receive full skylight (they actually receive much less), will equal $0.16\,B$. The mean apparent brightness will be equal to $\frac{1}{2}(0.8 + 0.16)\,B = 0.48\,B$. The mean apparent reflection factor would then equal $\frac{0.48\,B}{5\,B}$, or 0.1 approximately. Thus it is seen, with these assumptions, the apparent reflection factor of the grass plot is about 0.1 notwithstanding the fact that the actual reflection factor of an individual blade was assumed to be 0.16.

It will also be noted that the trapping of light by virtue of the "depth" of such a surface was not considered in the foregoing analysis, and also that the blades of grass are oriented in all directions. Both of these influences would tend materially to reduce the value of reflection factor. Woods are strikingly dark owing to these reasons and on the whole these considerations prepare us to expect a surprisingly low mean apparent reflection factor of the earth. Bare earth which has been packed by rain and baked by the sun is conspicuously brighter than the same soil freshly plowed even after the latter has dried. It is commonly noted that wet dirt is darker than the same dirt when dry. For example, a piece of blotting paper which, when dry, reflected 0.74 of the incident light reflected only 0.54 of the incident light when wet.

The brightnesses of all earth areas do not vary much with the angle at which they are viewed. This is not the case with the surface of water, as will be noted later. Owing to haze a photometer will generally yield greater values for very oblique measurements than for vertical measurements. In making the thousands of observations upon which Table I is based it is interesting to note that an overwhelming majority of the values of apparent reflection factor as obtained indiscriminately over the landscape are included within a range from 3.5 to 10 per cent.

Many attempts were made to obtain the average relative brightness or reflection factor of landscapes as a whole by measuring the ratio of the illumination of the lower side of a horizontal surface to

that of the upper side. This measurement was anticipated in designing the photometer, but inasmuch as it was necessary with this small instrument to lean out of the airplane, the latter intercepted some of the light reflected by the earth. Precautions were taken to reduce errors to a minimum and the values obtained were considered sufficiently accurate for their intended purpose. They were included well within the range from 5 to 10 per cent. The mean value of the reflection factor of the landscapes as a whole obtained in this manner was slightly greater than the mean obtained by weighing the values in Table I, but it will be noted that the latter are for perpendicular viewing and that the brightness of the landscape generally increases for oblique viewing owing to haze, to distant clouds, and to increased specular reflection from water. It is a simple matter to obtain the ratio of upward to downward light from an airplane by placing a ground opal glass at each end of a long vertical tube and determining the brightness of one in terms of the other by means of an instrument between them. This was done, but inasmuch as the mean reflection factor of the earth was of little importance in solving the problems primarily in mind, few readings were made. The values of the four types of areas were much more adaptable to the solution of the primary problem.

CLOUDS

Some measurements were made in the course of these investigations in order to establish the magnitude of the relative brightness or apparent reflection factor of clouds. These values vary considerably with the character of the cloud, the lighting conditions, etc. The apparent reflection factor of cumulus clouds receiving full skylight and sunlight as measured vertically above them on various days varied with the density of the clouds. The mean values of the upper surface of clouds, determined on different days, were found to be as follows:

Very dense clouds of extensive area and great depth	0.78
Dense clouds, quite opaque	.62
Dense clouds, quite opaque	.59
Dense clouds, quite opaque	.55
Dense clouds, nearly opaque	.44
Thin clouds	.40
Thin clouds	.36

One day when the sky was very heavily overcast with the exception of a small rift which admitted the airplane with safety to the upper region, a great many measurements of the brightness of the upper surface of the clouds were made. These observations were made near the sunny side of the hole in the extended massive layer of clouds and the values of relative brightness obtained varied from 0.7 to 0.8 with a mean of 0.78. It was somewhat surprising to obtain these high values, but they are perhaps accounted for by mass effect. Inasmuch as the upper side of a cumulus cloud was found to be six times as bright as the lower side we must conclude that there is a "mass" effect unless we choose to conclude that there is a tremendous absorption. In other words, great masses of clouds can be relatively much brighter than a purely diffusing, non-absorbing medium in a thin layer which theoretically would possess a transmission factor equal to 0.5 and a reflection factor of the same value. Of course white paint owes its high reflection factor chiefly to body or "mass" effect, but the depth to which the light penetrates is very small. The sunlit surfaces of dense clouds as viewed from below were found to be commonly 5 to 10 times the brightness of the adjacent patches of clear blue sky.

HAZE

The term "haze" is rather indefinite, for it may be taken to mean any atmospheric condition which increases the brightness of the sky (with the exception of definite clouds) or which intercepts light in any manner. There appear to be two general types of haze. One is the smoke and dust haze; the other is similar in appearance to very thin diffused cloud and is ofttimes perceptible at high altitudes beyond which the "dust" haze is not ordinarily found. The writer is not familiar with any definite distinction in nomenclature between these different types of haze but will adopt the term "earth" haze for the absorbing smoke or dust haze in order to distinguish it from the vapor haze, which is less absorbing and which apparently extends to much higher altitudes than the former. This haze will be termed "cloud" haze for the present, but with the understanding that no definite cloud shapes are apparent. Probably the classification could be carried farther to the higher regions of the atmos-

phere which account for the deep blue sky beyond the heights attained by cirrus clouds, by supplying a distinct name for this scattering of light at extremely high altitudes. During the progress of these investigations the earth haze extended ordinarily to an altitude of about a mile, although on a few occasions it extended to altitudes as high as 11,000 feet. This haze is somewhat absorbing, as is readily seen on viewing the earth obliquely, and is also fairly luminous. On some days at altitudes of 10,000 to 15,000 feet the sky appeared quite bright and unsaturated in color. As already stated, this upper haze often appears to be of the nature of a very thin, uniform, diffused cloud.

No very satisfactory measurements of the luminosity of this low-lying earth haze were obtained from above. Such an experiment requires a large, uniform, horizontal area at the earth's surface. However, a deep channel in the Chesapeake Bay was selected and measurements of its brightness were made as the airplane ascended over this area from 50 to 10,000 feet. Unfortunately the haze was unusually thin and little effect was detected. The greatest increase in brightness that could possibly be accepted from the brightness measurements was 10 per cent. Some measurements of the total illumination upon the upper surface of a ground opal glass at various altitudes were made without obtaining differences greater than the possible experimental errors in this case. This day was strikingly clear and cloudless.

The earth haze which absorbs as well as scatters light decreases the brightness of the earth, as viewed from high altitudes. Scattered light renders the haze luminous as viewed from above, which brightness tends to offset the decrease in the brightness of the earth. Viewing the matter in this manner it appears possible that the brightness of the earth plus the haze as viewed vertically does not vary appreciably for altitudes as high as 10,000 feet when there are no clouds or conspicuous vapor haze.

On another cloudless day similar measurements were made of the brightness of water at altitudes of 20,000 feet. The experiment was not as definitely planned as the preceding one, so that the measurements were less reliable; however, it appears possible that the brightness of the water areas increased as much as 25 to 50 per cent.

The presence and the effect of the earth haze, which ordinarily possesses a fairly definite upper surface, is easily apparent to the eye. The "false" horizon due to this dust haze is clearly visible, often with a few cumulus clouds partially submerged in it like icebergs. Many of the effects of air currents, the formations of clouds, the operation of winds and storms, can be viewed and studied so easily from the airplane capable of attaining high altitudes that it appears certain the airplane will be utilized by the scientist in studying these and many other problems.

SKY

During this and previous work many measurements of the sky brightness were made. The brightness of the sky generally varies from 0.5 to 2 lamberts. One lambert is a fair average as determined at the earth's surface. Very deep blue sky and dark overcast skies are approximately of the same brightness. The brighter skies are those that are very hazy or fairly heavily overcast. The brightest sky is that which is thinly overcast. On cloudless days the sky contributes from one-half to one-tenth of the total light, reaching a horizontal surface at noon. The larger value is usually encountered in winter, when the altitude of the sun is less than in summer. Measurements of this value should be for the same altitude of the sun in order to be strictly comparable. It is obvious that as the haze increases, the ratio of skylight to total light on the horizontal surfaces increases.

At low altitudes the sky is rarely so clear that only one-tenth of the total light which reaches the earth comes from the sky. Many measurements of this ratio were made at various altitudes with the obvious results. It is interesting to note in one case that at an altitude of 20,000 feet only 5 per cent of the total downward light was skylight notwithstanding the presence of a few thin cirrus clouds at altitudes of about forty-five degrees. The sun was at an altitude of perhaps fifty degrees, so that when corrected to zenith sun the value would be nearer 4 per cent. On this day at the earth's surface the sky contributed about 20 per cent of the total light. In other words, the sky was not more than one-fourth as bright when viewed from the high altitude as when viewed from

the earth's surface on this particular day. It was interesting to note how strikingly dark the sky appeared and how extremely harsh the lighting effect. That is, the shadows were noticeably very dark, especially if one is used to such observations.

On a very hazy but cloudless day measurements were made of the brightness of the zenith sky at altitudes from 0 to 15,000 feet. At the high altitudes the sky was extremely free from "vapor" haze and was very dark. In Fig. 3 are shown the results obtained in the

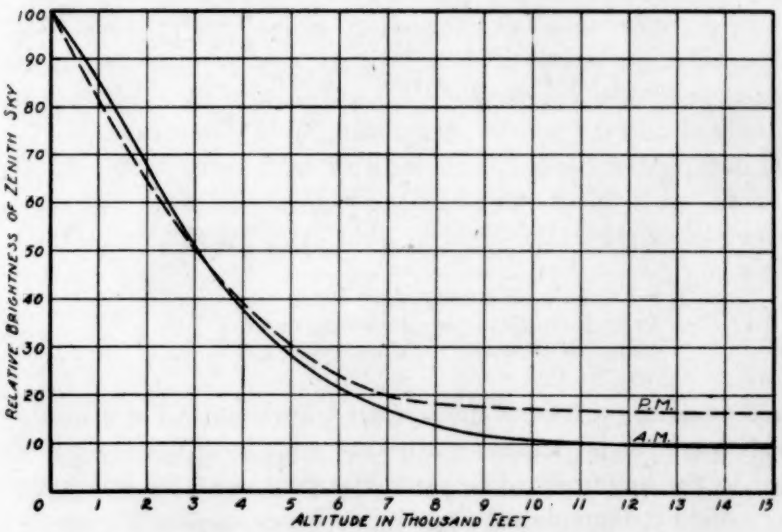

FIG. 3.—Relation of sky-brightness and altitude on a cloudless hazy day

morning and afternoon respectively. The region where the curve flattens out indicates the altitude where the earth haze ends. These regions are seen to be at about 10,000 feet and 7,000 feet, respectively, for morning and afternoon on this particular day, which was an exceptionally hazy one. Ordinarily there would not be such a great difference between the sky brightness as viewed from the earth's surface and from high altitudes, although the difference is usually considerable.

Although the brightness of clear sky is quite variable, it is of interest to know the brightnesses of various areas in terms of the sky brightness B, but these values should not be confused with

those values of "relative brightness" which are in reality "apparent reflection factors." The mean values of many observations obtained on various days for various areas are presented in Table II,

TABLE II

RELATIVE BRIGHTNESSES OF VARIOUS TYPES OF AREAS IN TERMS OF CLEAR BLUE SKY AS DETERMINED USUALLY AT LOW ALTITUDES
(Zenith sky = B)

Woods	Water	Fields	Plowed or Barren Land
0.44 B	0.53 B	1.02 B	.....
.36	.....	.67	.....
.45	.58	.86	.....
.50	.73	.....	.....
.40	.80	.77	1.08 B
.53	.91	.93	1.19
	.72	.62	.....
Mean 0.45 B	0.71 B	0.81 B	1.14 B

Landscape covered with hazy clouds, 3.5 B
Upper sunlit sides of cumulus clouds, 5–10 B
Under side of massive cumulus clouds, 1.4 B

the sky brightness B being unity in arbitrary units. The measurements were made simultaneously and usually at low altitudes. That is, the brightness of an earth area relative to the zenith sky was made by comparing the two brightnesses simultaneously.

The values for clouds and for water are discussed in other sections under their respective headings. It is interesting to note that the data in Table II were obtained independently from those in Table I with an instrument of a different type, and in terms of the blue sky which varies from hour to hour and from day to day. However, when reduced to the same value for water they compare thus:

	Fields	Barren Lands	Woods	Water
Table I.....	6.8	13.0	4.3	6.8
Table II.....	7.7	10.9	4.3	6.8

This is an excellent check when all aspects are considered.

A number of determinations of the portion of the total light contributed by the sky on clear days were made at the earth's surface during this period. These averaged about 0.16. In other words, the sky contributed about one-sixth and the sun about five-sixths of the total light on the earth's surface for the clear days of this period and in this locality. However, the values varied considerably. Let us assume the mean sky brightness equal to B and the brightness of a perfectly reflecting and diffusing horizontal surface due to sunlight equal to $5\,B$. By weighing the values of earth brightness in terms of the mean sky brightness (which is greater than the brightness of the zenith sky) let us assume the mean brightness of the earth equal to $0.4\,B$. This value is lower than those in Table II because it is not uncommon for the sky near the horizon to be several times brighter than the zenith sky brightness. The values in Table II would be considerably smaller if they were in terms of average sky brightness. Based on these facts and assumptions,

$$R\,B\,(5+1)=0.4\,B$$

$$R=\frac{0.4}{6}=0.066$$

where R is the mean apparent reflection factor of the earth. This is approximately the weighted mean arrived at by measurement. Inasmuch as the assumptions cannot be far from correct this serves as an approximate check for the data in Table I.

Some measurements at altitudes of 10,000 to 20,000 feet on a clear day yielded the following brightnesses in terms of the zenith sky brightness B viewed from these altitudes.

Fields, etc.	$1.1\,B$
Water	$1.4\,B$
Landscape covered with thin clouds	$3.5\,B$
Dense sunlit clouds	$11.0\,B$

A mean of a number of measurements at 18,000 feet indicated that the lower horizontal surface of a perfectly reflecting and diffusing opaque object (reflection factor = 1) due to light reflected upward by the earth was 1.7 times the brightness of the zenith sky as observed at that altitude. In accordance with the preceding

computation, taking account of the fact already stated that at this altitude the sky was of uniform brightness and contributed 0.05 of the total light,

$$R\,B\,(19+1)=1.7\,B.$$

Therefore R, the mean apparent reflection factor of the earth, would equal 0.085. Considering that the earth haze is responsible for some of the light which reaches the lower horizontal surface of the theoretical perfectly reflecting and diffusing object, the value 0.085 appears to check well with preceding values.

Various observations of the brightness of the lower horizontal surface of a perfectly reflecting and diffusing object were determined in terms of the brightness B of the zenith sky. A mean of observations of this value made at altitudes from 7000 to 13,000 feet with a few clouds at 5000 feet was found to be $0.9\,B$. A mean of observations of this value made on another day at an altitude of 7000 feet with some clouds at 5000 feet (2000 feet below) was found to be $1.6\,B$. A mean of observations of this value at altitudes from 10,000 to 18,000 feet on a fairly clear day was found to be $1.7\,B$. In these cases there were no clouds directly below. The differences in the values presented above are due doubtless to different brightnesses of the sky, to clouds, and to various other lighting conditions. From the mean of the three mean values and assuming an average apparent reflection factor of the earth equal to 0.085 (including the contribution of the haze) we compute that the upper horizontal surface of a perfectly white, opaque diffusing object would appear about 12 times brighter than the lower surface of such an object above the clouds and haze on a day representing a mean of these three days.

From the data in Table I, assuming a mean apparent reflection factor of the earth equal to 0.07, the upper surface would be about 14 times brighter than the lower surface when the opaque object is a few thousand feet above the earth. Under these conditions the brightness of the lower side of the opaque object would equal the brightness of the sky (assuming it to be uniformly bright) on a clear day if the sky contributed one-fourth of the total light. This latter condition is perhaps never obtained at very low altitudes. At low altitudes up to 5000 feet the lower horizontal surface of a perfectly reflecting and diffusing object is usually only a fraction

of the brightness of the zenith sky. It should be noted that this fraction increases with altitude due chiefly to the decrease in the sky brightness.

It is interesting to note that Kimball and Thiessen[1] obtain the following ratios of illumination on a horizontal surface for clear and smoky atmosphere:

$$\frac{\text{Total illumination on smoky day}}{\text{Total illumination on clear day}} = 0.64$$

$$\frac{\text{Direct sunlight on smoky day}}{\text{Direct sunlight on clear day}} = 0.48$$

$$\frac{\text{Skylight on smoky day}}{\text{Skylight on clear day}} = 1.64$$

This shows how the sky brightness is increased by the presence of reflecting particles.

WATER

As seen in Table I the mean relative brightness or apparent reflection factor of inland waters as obtained by measurements perpendicular to its surface was 0.068. Practically all the values obtained on different days were within the range from 0.05 to 0.1 and most of them were close to the mean value. If the surface of water is perfectly smooth it will reflect at its surface about 0.02 of perpendicularly incident light; in other words, its brightness will be 0.02 of the brightness of the sky or clouds vertically above. This reflection factor varies only slightly with increasing angle of incidence until about fifty degrees is reached. It begins to increase more rapidly from this angle until obviously at an angle of incidence of ninety degrees it reflects all the light.

Owing to the fact that the brightness of an area of water, viewed from above, is due partially to specular reflection, its brightness will depend partially upon the brightness of the objects whose images are reflected. These objects are blue sky, haze, clouds, and the sun; however, the latter will be neglected in this discussion because it is a very special case. It will be seen in Table I that the mean value of the relative brightness of water determined perpendicularly to its surface is fairly constant on different days. In fact this value varies surprisingly little for deep water which would be considered

[1] *Monthly Weather Review*, **45**, 205, 1917.

fairly clear. An increased brightness is noted for shallow water and for water which holds bubbles or fine particles of solid matter in suspension. For example, measurements on the James River about 30 miles from its mouth gave values of 0.10 for shallow water near the shore and 0.075 in midstream. The latter value was very constant for its deep channel. Near Richmond, where the river was very yellow, due to suspended particles of earth, the relative brightness rose to 0.20, which is very unusual.

The foregoing values were obtained for inland waters within a radius of 150 miles of Hampton, Va., and included Chesapeake Bay, Hampton Roads, and various rivers. In order to obtain values of apparent reflection factor or relative brightness of deep ocean waters far from shore, a trip over the Atlantic was arranged. A series of measurements was made during a six-hour flight over Hampton Roads, Chesapeake Bay, through the Virginia Capes and 50 miles to sea, at altitudes averaging about 1000 feet. As usual the measurements were made approximately at normal incidence with the surface of the water. The mean values of apparent reflection factor follow:

	Percentage
Shallow water in Hampton Roads	7.4
Deep channel in Chesapeake Bay	4.9
Atlantic Ocean 10 to 50 miles from land	3.5

These represent means of many observations. The lowest value obtained for deep water far from land was about 2 per cent and the highest 5.6 per cent. Most of the values were confined to a range from 3 to 5 per cent. Of course the brightness increased for oblique angles of view and varied with the position of the sun.

The difference in the hues of shallow turbid water and deep clear water is quite apparent. When the water is extremely shallow the hue partakes partially of the color of the bottom. In moderately deep inland waters which are in motion and consequently more or less turbid the water appeared a dirty greenish hue. The deep clear ocean water is by comparison a much purer blue-green hue.

In Table III are presented some values of the brightness of inland waters in terms of the brightness of the zenith sky. These measurements were made vertically and simultaneously.

Similar measurements were made on one day for deep ocean water which resulted in values from 0.16 to 1.0 times the brightness of the zenith sky. This extreme range was due largely to the changing brightness of the sky on this day. Early in the day the sky was a clear deep blue, but gradually became more hazy and finally was covered with thin cirrus clouds.

TABLE III

RELATIVE BRIGHTNESSES OF INLAND WATERS ON VARIOUS DAYS VIEWED PERPENDICULARLY, IN TERMS OF THE BRIGHTNESS OF THE ZENITH SKY *B* AS DETERMINED FROM ALTITUDES UP TO 20,000 FEET

4:30 P.M.	Sunny day, large clouds	0.58 *B*
10:30 A.M.	Clear day	.63
10:00 A.M.	Clear day, some clouds	.80
9:30 A.M.	Clear day, some clouds	.87
2:00 P.M.	Clear day	.99
1:30 P.M.	Clear day, deep blue sky	.79
1:30 P.M.	Clear day, some clouds, high altitude	1.40
Mean		0.87 *B*

From the position of the observer directly overhead on clear days the surface of the water reflected only images of portions of the blue sky regardless of the character of the waves. Inasmuch as surface reflection could account for only a small portion of the mean value in Table III, it is obvious that most of the brightness of water is due to light diffused within it. This is well verified by the presence of cloud shadows commonly seen upon the water. These are especially conspicuous in aërial photographs of water areas when cumulus clouds are present to cast deep shadows and are strikingly evident when the water is viewed at an oblique angle from high altitudes.

In this connection it is interesting to note that some small lakes and ponds are strikingly dark amid their surroundings, sometimes appearing almost black. In fact these are often the darkest areas in a landscape. If a pool is quiescent, clear, and deep, with a porous bottom of low reflection factor such as is often the case, practically the entire brightness is that due to specular reflection from the surface. A smooth surface of water reflects only about 0.02 of normally incident light, and as viewed vertically its brightness due to

surface reflection is only 0.02 that of the zenith sky. On a clear day, if the sky contributes only one-fifth of the total light, the brightness of such a quiet, clear, deep pool (assuming a sky of uniform brightness) would be only 0.004 or 0.4 per cent of a perfectly white surface. In other words, under these conditions it would appear of about the same brightness as a black velvet fabric. As seen in Table I, inland water is about as bright as grassy plots of ground, which shows that most of the brightness of water is due to its turbidity, that is, to the light diffused within it. On the foregoing assumptions about 90 per cent of the brightness of water viewed perpendicularly is due to light diffused within it.

Some values of the brightness of areas of water viewed at forty-five degrees in terms of the brightness of the zenith sky varied from 1.5 to 3.0. The mean of the few values obtained is about 2.2. These values varied considerably at any given time, depending upon the direction of observation with respect to the sun. This is one of the reasons which led to making practically all the observations perpendicularly to the earth's surface.

In such measurements as have been presented in this paper, the variability in natural lighting, in sky brightness, etc., causes considerable uncertainty in some cases. The values of reflection factor of earth areas, water, and clouds represent closely the true values for the summer landscapes. The values involving the brightness of the sky indicate no more than limits or approximate magnitudes. Owing to the variation in sky brightness an exhaustive research extending over a long period would be necessary in order to establish mean values.

NELA RESEARCH LABORATORY
NATIONAL LAMP WORKS OF GENERAL ELECTRIC CO.
NELA PARK, CLEVELAND, OHIO
November 1918

Reviews

The Binary Stars. By Robert Grant Aitken. New York: Douglas C. McMurtrie, 1918. 8vo. Pp. xiv+316, plates 5, figs. 12. $3.15.

This volume is an important addition to our literature on stellar astronomy. It is especially welcome because written by a man who is in a better situation than anyone else for treating the subject. His personal contributions to the field covered by the book are very important. His sphere of activity has been for years one where the newer branches of double-star astronomy, as developed by the use of the spectroscope, have been cultivated successfully and intensely. Through these circumstances the different angles of the subject have been given an adequate attention and the whole field is covered in a well-balanced way.

Some forty years ago the *Handbook of Double Stars*, by Crossley, Gledhill, and Wilson (London, 1879), could be considered as embodying our essential knowledge in the field of binary stars at that time. Nothing can make us realize better the progress made between the publication of this book and the present time than to compare the *Handbook* with Mr. Aitken's volume. How the field has been enlarged! What unexpected developments have come to side-lines of the subject in the meantime!

The first two chapters of the book outline the historical evolution of the study of double stars. Beginning with the first isolated discoveries, it is followed by the exploring work of the Herschels and Struves and the standardizing of the visual methods of observing. Next we come to the startling discoveries by Burnham in the seventies, the extension of the field of the binaries through stars revealed by the spectroscope and those that have become known as eclipsing variables until, with the systematic survey of the northern sky at the Lick Observatory, we arrive at the present epoch.

The author considers next the technical side of the observations of visual binaries with the micrometer. The instrument used by him is described in detail and also the methods of measuring that he has practiced. He strongly advocates the advantages of having only two parallel

wires, a fixed and a movable one, in the field. In the reviewer's opinion Mr. Aitken is going too far when he adds, on page 41: " . . . . and even for comets, asteroids, satellites, or other forms of micrometric work I regard it as superior to the more complicated forms. . . . . " In the case of comets or asteroids it is generally not possible to connect those moving objects with a catalogue star by using direct differential measures. Generally they have to be referred to some fainter star whose position is found by connecting it to a known star by transits. Here the observations with only two wires are decidedly too slow as compared with what is done when the micrometer has three or five transverse wires so that both co-ordinates can be obtained simultaneously with about the same accuracy; this disadvantage cannot be overlooked when we bear in mind what the author says on page 55: "An hour in the dome on a good night is more valuable than half a dozen hours at the desk in daylight." Also in the determination of the parallel a long transverse wire gives quicker and more accurate results than the short parallel wires. In turning off the illumination of the transverse wires, we can use the micrometer for double-star work as if there were only two parallel wires.

As to the determination of the value of the revolution of the micrometer screw (page 43) we would make this comment: One of the methods advised consists in taking transits of circumpolar stars over the movable wires in successive positions in the field. "Theoretically a correction for refraction should also be introduced, but if all of the measures are made near the meridian this correction will rarely be appreciable." That this correction is quite appreciable even *on the meridian* can be seen from the following table, corresponding to a latitude of 42°.

Declination	Correction in Right Ascension
+20°	1 part in 4200
40	1 part in 3600
60	1 part in 2300
80	1 part in 1900

The corrections are fully of the same order as the corrections in declination as we find them on page 47 in Mr. Barnard's measures of Electra and Celaeno taken about two hours past the meridian: 0.020 in 65.57 Rev. or 1 in 3300.

This same chapter condenses a great amount of useful information about the work with the micrometer. Every new student in this field should carefully read the sound remarks about the choice of an observing

program, in which Mr. Aitken shows himself a worthy disciple of Mr. Burnham, to whom he dedicates this book.

In chapter iv we find the problem of computing double-star orbits clearly set forth. No original views are presented but the more generally used methods are made very plain. Graphical methods are considered more especially in detail and a numerical example of an application of the Glasenapp-Kowalsky method is given as an illustration of the determination of a preliminary orbit, that of the rapid binary A 88 with a period of but 12 years. When the observing material justifies it, a further approximation is obtained by differential corrections. No mention is made here of Doberck, who does away with the complicated analytical expression for the differential coefficients but finds these by changing arbitrarily the successive elliptic elements and computing the representation of the normal observations. He computed tables for this many years ago giving the values of the true anomalies and radii vectores in function of the mean anomalies for different values of *e;* these tables, which are much handier than Åstrand's, giving eccentric anomalies to which the author refers (p. 74), have, however, been superseded recently by the so-called Allegheny tables whose importance cannot be overestimated by the double-star computers. Since these are at hand no computer should go any longer to the trouble of calculating the analytical formulae for the differential coefficients.

Further paragraphs are devoted to the special cases in the orbits: the case of an orbit inclined at 90° so that the apparent oscillating motion is rectilinear, with an original contribution of F. R. Moulton (p. 96) on the subject; the case of an apsidal line coinciding with the nodal line, as treated by H. N. Russell (p. 99); the systems where only one component is visible, and, finally, the case of rectilinear motion.

The next chapter, written by J. H. Moore of the Lick Observatory, gives with regard to the spectroscopic binaries what the first part of the book has brought in regard to the visual ones. To find these two fields brought so closely together and treated as far as possible along parallel lines, will be gratifying to most readers. After explaining the principle of the Doppler-Fizeau method, Mr. Moore discusses the instrumental devices that are necessary for bringing the spectrograms to the high standard that they have reached in a relatively small number of years, and also the methods used in measuring and reducing the spectra. A plate giving some typical spectra is found on p. 114, but some readers will miss an indication of wave-length of the principal lines.

Chapter vi corresponds for the spectroscopic binaries to chapter iv for the visual ones. Mr. Aitken gives here the methods generally used in computing the orbits and the many new ones that have been proposed in recent years. The principal method is again illustrated by a numerical application to κ Velorum.

The stars where the binary character is found from periodic variations of the apparent brightness are dealt with in chapter vii. Mr. Aitken follows closely the theoretical investigation of the problem as it was worked out by H. N. Russell and applied with great success by H. Shapley to the large number of stars already found in this class. The principal numerical tables used in Russell's work are reprinted here so that it is not necessary to refer to the original memoir for using the method. Here also we find a numerical illustration, that of the well-known Algol star, W Delphini. No description is given of the instruments or methods used in observing these eclipsing binaries. This subject would cover the whole field of astronomical photometry for which there is already an abundant literature.

With chapter viii we come to the second part of the book and the discussion of the facts that have been collected in the various ways described in the first part. The valuable list of orbits for spectroscopic binaries (137) as well as visual binaries (87) tabulated at the end of the volume forms the basis of this investigation. The close relations between the length of the periods, the size of the systems, and the degree of ellipticity of the orbits are clearly brought out. The preference of the binaries for certain spectral classes is another striking feature: the visual ones clustering in class G, the spectroscopic just as decidedly in classes B and A, and also the increasing masses of the systems as we come to earlier types.

Next the author considers individually a selection of stars that are interesting from different points of view. The history of such stars as α Centauri, Sirius, Polaris, Algol, etc., is always full of interest and the author has made this chapter ix one of the most attractive for reading.

In chapter x, as a result of the systematic survey of the northern sky, recently completed by the author, he discusses in a statistical way the number and distribution of the visual doubles. The main conclusion, that at least 1 in every 18 of the stars north of the equator that are at least as bright as $9^{M}.0$ in the BD is a close double in the 36-inch telescope, has already been announced previously by the author. It is supplemented by the fact that close visual double stars are relatively more

abundant in the Milky Way. Many other facts in connection with the spectral types are also discussed here.

The final chapter is devoted to the cosmogonic theories that have been proposed for explaining the formation of so many binary stars: the capture theory, the fission theory, and the theory of independent nuclei. Each of these hypotheses explains a certain number of the facts that have been established by the study of double stars. But none of them is entirely satisfactory when all the facts are considered. It is very difficult to decide yet in favor of one of these. In the opinion of the author, the fission theory would perhaps lead to the most natural explanation of the observed facts.

To co-ordinate in one comparatively small volume such a large variety of lines of research was not an easy task. Mr. Aitken must be congratulated for doing it so successfully and for having presented the subject in such an attractive way.

G. VAN BIESBROECK

YERKES OBSERVATORY

EDWARD CHARLES PICKERING

With the deepest regret we record the death on February 3, 1919, of EDWARD CHARLES PICKERING, in the seventy-third year of his life and the forty-second year of his imperishable service to science as director of the Harvard Observatory.

He had been associated with this Journal as Collaborator since its foundation.

A biographical sketch will appear in a later number.